FLO11 gene length and transcriptional level affect biofilm-forming ability of wild flor strains of *Saccharomyces cerevisiae*

Giacomo Zara,^{1†} Severino Zara,^{1†} Claudia Pinna,¹ Salvatore Marceddu² and Marilena Budroni¹

Correspondence
M. Budroni
mbudroni@uniss.it

¹Dipartimento di Scienze Ambientali Agrarie e Biotecnologie Agro-alimentari (DISAABA), University of Sassari, Viale Italia 39, 07100 Sassari, Italy

²Istituto di Scienze delle Produzioni Alimentari (ISPA CNR Sassari), Via dei Mille 48, 07100 Sassari, Italy

In *Saccharomyces cerevisiae*, *FLO11* encodes an adhesin that is associated with different phenotypes, such as adherence to solid surfaces, hydrophobicity, mat and air–liquid biofilm formation. In the present study, we analysed *FLO11* allelic polymorphisms and *FLO11*-associated phenotypes of 20 flor strains. We identified 13 alleles of different lengths, varying from 3.0 to 6.1 kb, thus demonstrating that *FLO11* is highly polymorphic. Two alleles of 3.1 and 5.0 kb were cloned into strain BY4742 to compare the *FLO11*-associated phenotypes in the same genetic background. We show that there is a significant correlation between biofilm-forming ability and *FLO11* length both in different and in the same genetic backgrounds. Moreover, we propose a multiple regression model that allows prediction of air–liquid biofilm-forming ability on the basis of transcription levels and lengths of *FLO11* alleles in a population of *S. cerevisiae* flor strains. Considering that transcriptional differences are only partially explained by the differences in the promoter sequences, our results are consistent with the hypothesis that *FLO11* transcription levels are strongly influenced by genetic background and affect biofilm-forming ability.

Received 26 February 2009

Revised 22 July 2009

Accepted 28 August 2009

INTRODUCTION

Saccharomyces cerevisiae adapts to environmental changes by remodulating the composition of the cell wall. Under nutrient starvation, *S. cerevisiae* cells can interact with each other, and with biotic and abiotic surfaces (Verstrepen & Klis, 2006). The resulting cell aggregation is mediated by adhesins, the *FLO* gene family proteins that confer phenotypic diversity and great plasticity to the cell wall. Among these *FLO* genes, *FLO11/MUC1* is involved in adhesion to plastic surfaces and invasive growth on agar. *FLO11* encodes a GPI-anchored glycoprotein that is rich in serine and threonine, and that mediates homotypic adhesion between cells (Douglas *et al.*, 2007). *FLO11* is also involved in substrate–cell interactions, such as biofilm formation on plastic and liquid surfaces (Reynolds & Fink, 2001; Zara *et al.*, 2005). $\Delta flo11$ mutants cannot form biofilms, allowing for the determination of the role of *FLO11* in biofilm formation on liquid surfaces (Ishigami *et al.*, 2004; Zara *et al.*, 2005; Fidalgo *et al.*, 2006).

As with other cell-wall-related genes, the *FLO11* ORF is characterized by intragenic repeat sequences (Mannazzu *et al.*, 2002; Marinangeli *et al.*, 2004; Verstrepen *et al.*, 2005). In *S. cerevisiae*, polymorphism of protein size is quite frequent and is often associated with insertion/deletion (indel) events in the DNA that result in repeated regions that are rich in serines and threonines (Bowen *et al.*, 2005; Verstrepen *et al.*, 2005; Bowen & Wheals, 2006). Recent evidence has suggested that this polymorphism is likely to be under selection, as differently sized orthologues show differing levels of mutual attraction (Verstrepen *et al.*, 2005). The *FLO11* gene carries one of the largest promoter regions of the yeast genome. The *FLO11* promoter is at least 2,800 bp (Lo & Dranginis, 1998) and it is complex: it consists of four upstream activation sequences and at least nine upstream repression sequences (Pan & Heitman, 1999; Rupp *et al.*, 1999; Pan & Heitman, 2002; Kuchin *et al.*, 2002), the activities of which depend on growth stage and nutritional conditions (Rupp *et al.*, 1999). Some of these regulatory regions are targets for the MAP kinase pathway, the adenylyl cyclase/cAMP cascade and the Snf1-Nrg1/Nrg2 pathway (Kuchin *et al.*, 2002). Fidalgo *et al.* (2006) identified a 111 nt deletion within an upstream repression sequence of the *FLO11* promoter of a flor strain which increased gene expression. Apart from being targeted by

[†]These authors contributed equally to this work.

The GenBank/EMBL/DDBJ accession numbers for the *FLO11* promoter sequences of strains A9, V23, M23 and V80 are FJ526742, FJ526743, FJ526744 and FJ526745, respectively.

these different signalling cascades, *FLO11* is also under epigenetic control and chromatin remodelling complexes have fundamental roles (Frieman & Cormack, 2004; Halme *et al.*, 2004; Barrales *et al.*, 2008).

In the present study, we have evaluated the gene polymorphism of *FLO11* in 20 wild flor strains of *S. cerevisiae* isolated from three wine-producing areas of Sardinia, Italy. We show that in flor strains, the propensity for biofilm formation depends on the length and transcription level of *FLO11* alleles, and that *FLO11* transcriptional levels are strongly influenced by the genetic background.

METHODS

Yeast strains and media. The yeast strains used in this study are listed in Table 1. *S. cerevisiae* flor strains from the DISAABA culture collection (University of Sassari) were subdivided into three different clades, as suggested by previous results from PFGE, internal transcribed spacer (ITS)-RFLP, non-transcribed spacer (NTS)-RFLP and mtDNA-RFLP analyses (Farris *et al.*, 2002). Strain BY4742 is isogenic to S288C, which bears a nonsense mutation in the *FLO8* gene that abolishes *FLO11* expression (Liu *et al.*, 1996). All of the strains were cultivated on YPD (1 % yeast extract, 2 % peptone, 2 % glucose) and synthetic complete (SC) medium (0.67 % yeast nitrogen base without amino acids, 2 % glucose, supplemented with bases and amino acids at standard concentrations) at 30 °C (Sherman *et al.*, 1986). The flor medium was Yeast Nitrogen Base (YNB), containing 4 % ethanol as sole carbon source, which was supplemented where necessary with bases and amino acids at standard concentrations (Zara *et al.*, 2005).

Molecular methods. Genomic DNA isolation was performed according to Burke *et al.* (2000). PCRs for *FLO11* gene amplification were performed using the MUC1up and MUC1lo primers (Table 2), with 100 ng genomic DNA, 3 mM MgCl₂, 0.25 mM dNTP, 50 pmol primer Up, 50 pmol primer Lo, and 0.5 U *Taq* DNA polymerase (EuroClone) in a 50 µl volume. The temperature cycling parameters were as follows: an initial hold at 95 °C for 5 min; 40 cycles at 95 °C for 1 min, 52.9 °C for 50 s and 72 °C for 4 min; and a final elongation at 72 °C for 15 min.

The EMBOSS programs EQUICKTANDEM and ETANDEM (Rice *et al.*, 2000) were used to identify intragenic repeats in *FLO11* reference sequences in databases (Saccharomyces Genome Database,

www.yeastgenome.org; GenBank www.ncbi.nlm.nih.gov/Genbank). PCRs for tandem repeat domains in each *FLO11* allele were carried out using the primers Up776flo11 and Lo2973flo11, which were designed on the basis of the S288C sequence (Table 2). The temperature cycling parameters were as follows: an initial hold at 94 °C for 5 min; 30 cycles at 94 °C for 1 min, 60 °C for 50 s and 72 °C for 3 min; and a final elongation at 72 °C for 10 min. The amplification of *FLO11* intragenic repeats was followed by a high-resolution melting analysis which allows the estimation of the base composition of the PCR products (Stephens *et al.*, 2008).

FLO11 ORFs of A9 and M23 strains were cloned into the pYES2.1/V5-His-TOPO vector (Invitrogen), which contains a *GAL1* promoter and a *CYC1* transcriptional terminator downstream of the insertion site. Insertion was confirmed by PCR using the primers GAL1 Forward and V5 C-term (Invitrogen). The laboratory strain BY4742 was transformed with the two plasmids, yielding new strains FA9 and FM23, which contained the *FLO11* ORF of A9 and M23, respectively.

The *FLO11* promoters of the M23, V23, A9 and V80 strains were amplified using primers Up prom flo82 and Lo prom flo3022 (Table 2). The PCR products were directly sequenced at BMR Genomics, Padova, Italy, by Genome Sequencer FLX (Roche) using the primers listed in Table 2. Sequence alignments were performed using BLAST 2 (www.ncbi.nlm.nih.gov/blast/bl2seq/wblast2.cgi) and CLUSTAL W (www.ebi.ac.uk/Tools/clustalw2/index.html) for single and multiple alignments, respectively.

***FLO11*-associated phenotypes.** Formation of an air-liquid interfacial biofilm was performed as follows. The strains were grown in 2 ml YPD overnight at 30 °C in an incubator-shaker, recovered by centrifugation, washed once in sterile distilled water, and resuspended in 2 ml flor medium in borosilicate test tubes. Samples were incubated at 30 °C for 3–5 days under static conditions. The biofilms were harvested by aspiration with a vacuum pump and filtered using previously weighed 2 µm nitrocellulose filters. The filters were then oven-dried at 50 °C for 72 h and weighed again. The biofilm was expressed as the difference between the weights of the filters with the dried pellets and the original weights of the dried filters.

Yeast adherence to polystyrene, aqueous-hydrocarbon biphasic hydrophobicity and colony morphology were evaluated as described by Reynolds & Fink (2001). For determination of the adherence to polystyrene surfaces, the cells were grown overnight at 30 °C in SC supplemented with 2 % (w/v) glucose, washed once in sterile water, resuspended in SC supplemented with 0.1 % glucose, and brought to an OD₆₀₀ of 1.0. Ten repetitions of 100 µl aliquots were then transferred into the wells of 96-well polystyrene plates and the cell suspensions were incubated at 30 °C for 3 h. An equal volume of 1 %

Table 1. *Saccharomyces cerevisiae* strains used in this study

Strain	Description	Source or reference
V16, V19, V23, V30, V44, V75, V80	Wild flor strains of <i>S. cerevisiae</i>	Isolated from Vernaccia wine (DISAABA collection)
M12, M23, M25, M39, M46, M49, M66	Wild flor strains of <i>S. cerevisiae</i>	Isolated from Malvasia wine (DISAABA collection)
A9, A43, A28, A68, A51, A41	Wild flor strains of <i>S. cerevisiae</i>	Isolated from Arvisionadu wine (DISAABA collection)
BY4742	<i>MATα his3Δ1 leu2Δ0 lys2Δ0 ura3Δ0</i>	Brachmann <i>et al.</i> (1998)
FM23	BY4742 transformed with plasmid pYES2.1/V5-His-TOPO bearing <i>FLO11</i> gene from M23 wild strain	This study
FA9	BY4742 transformed with plasmid pYES2.1/V5-His-TOPO bearing <i>FLO11</i> gene from A9 wild strain	This study

Table 2. Primers used in this study

Primer	Sequence (5'–3')
MUC1up	AGGTTCAAATGGTGCCAAGA
MUC1lo	AGCCACGCTAGAAAGCAGAAG
FLO11Rtup	AGGTTCAAATGGTGCCAAGA
FLO11Rtlo	AGCCACGCTAGAAAGCAGAAG
ACT1 up	CGTTCCAATTACGCTGGTT
ACT1 lo	TCAGCAGTGGTGGAGAAAGA
Up776flo11	CAAGGTAGTGCCGCTCAATA
Lo2973flo11	CAGTAGTGCCAGTAGAAAAG
Up prom flo82	CCAACTAAATCTGAATAACAA
Lo prom flo3022	AAGCGAAAGGACCAATAAGC
Seq prom flo134	GCAATGATTATGTGGTAT
Seq prom flo584	AGCTGAAAAGTCCATCTA
Seq prom flo1071	ACATCTTTGCTCCCTTAC
Seq prom flo1538	GGTGAGATTTGTTTATG
Seq prom flo1837	AATGTCGCCCCAAGAGTT
Seq prom flo2022	TGCGACAGTGGCTTCAAA
Seq prom flo2549	TGTGGGTCATCTTTTATG

(w/v) crystal violet was then added to each well. After 30 min, the wells were washed with sterile water, and the adherence of the cells was quantified by solubilizing the retained crystal violet in 100 µl 10% (w/v) SDS and an equal volume of sterile water. After 30 min, 50 µl of these solutions was transferred to fresh polystyrene 96-well plates, and A_{570} and A_{590} were measured spectrophotometrically.

For determination of the aqueous-hydrocarbon biphasic hydrophobicity, the cells were grown overnight in SC supplemented with 2% (w/v) glucose, washed once with water, and resuspended in SC supplemented with 0.1% glucose to an OD_{600} of 0.5. After a stationary incubation of 3 h at 25 °C, the OD_{600} of the cultures was measured again. Then, 1.2 ml of the cultures was added to each of three borosilicate glass tubes and overlaid with 600 µl *n*-octane. After vortexing the tubes for 3 min, the phases were allowed to separate and the OD_{600} of the aqueous layers was measured. The differences between the OD_{600} of the aqueous phase before and after the addition of *n*-octane were used to determine the hydrophobicity.

Finally, colony morphology on soft agar was evaluated by inoculating the yeast strains onto YPD soft agar plates (0.3% agar) with a toothpick 1–2 days after the plates were poured. The plates were then wrapped with Parafilm and photographed after 10 days incubation at 25 °C.

Scanning electron microscopy (SEM). The biofilms were fixed in 2.5% glutaraldehyde in 0.1 M sodium cacodylate buffer for 90 min, then washed in the same buffer and post-fixed in 1% osmium tetroxide solution for 1 h. They were then dehydrated in a series of increasing acetone concentrations (20, 50, 70, 80, 95 and 100%), dried under CO_2 with a Polaron critical-point device, sputter-coated with gold/palladium in an Edwards S150A sputter coater, and finally examined in a Zeiss DSM 962 scanning electron microscope.

FLO11 transcription analysis. Transcription analysis was performed on yeast strains grown on YPD and on flor medium. For each sample, 1×10^8 cells were collected after 8 h incubation in YPD and after 48 h static incubation in flor medium just before a visible biofilm was formed. RNA extraction and reverse transcription were carried out using RNAqueous-4 PCR (Ambion) and the SuperScript First-Strand Synthesis System for RealTime-PCR kit (Invitrogen), respectively. The cDNA was quantified using an iCycler iQ RealTime-PCR detection system (Bio-Rad). The primers used were ACT1 up

and ACT1 lo for *ACT1* amplification, and FLO11RTup and FLO11RTlo for *FLO11* amplification (Table 2). The SYBR-Green I dye (Sigma-Aldrich) was used as the fluorescence system. The reaction conditions were: 1 cycle at 95 °C for 3 min; 50 cycles at 95 °C for 20 s, 60 °C for 30 s and 72 °C for 30 s; and a final elongation at 72 °C for 3 min. The *ACT1* gene was used as the housekeeping gene to normalize the expression data across the different samples. The relative expression values were determined using the formula proposed by Pfaffl (2001).

Statistical analyses. To test the null hypothesis that the data sampled came from a normally distributed population, the Shapiro–Wilk test was used. The significance of the differences among samples was assessed by using the *t*-test. Considering the lack of information about the population distributions underlying the phenotypic variables measured, the non-parametric Spearman rank correlation coefficient was calculated. Regression analysis was used to estimate the linear model $[y_i = f_1(x_{1i}) + f_2(x_{2i}) + \varepsilon]$. The length of the intragenic repeats (x_1) and the transcription levels (x_2) of *FLO11* were the predictor variables and the weight of the biofilm formed was the response variable (y). The validity of the linear model was assessed by checking the absence of multicollinearity among the predictor variables and model departures from: (i) linearity, by non-parametric smoothing techniques (smoothing splines); (ii) normality, by graphic analyses (normal quantile–quantile residual plots); (iii) homoscedasticity, by graphic analyses (residual vs predictor plots). All of the statistical tests and calculations were performed using the free software R (version 2.8.1) (R Development Core Team, 2008).

RESULTS

FLO11-associated phenotypes are highly variable

The 20 *S. cerevisiae* flor strains exhibited a high diversity in their *FLO11*-associated phenotypes (Table 3). Colony morphologies on soft agar were different both for the diameters of the mat formed (ranging from 4.43 to 7.73 cm) and for the presence/absence of structures, such as cables and radial spokes emanating from the central hub (Fig. 1). The hydrophobicity values ranged from 61.26 to 98.95%, while the adhesion values ranged from 1.04 to 2.44 (A_{570}). Interesting differences were also seen for the weights (ranging from 0.12 to 0.32 mg), and the morphologies of the biofilm formed (Fig. 2). Also, for the first time in an air–liquid biofilm, SEM revealed the presence of an abundant extracellular matrix of unknown composition in the velum formed by some strains (Fig. 3).

FLO11 is highly polymorphic

PCR analysis of the 20 flor strains revealed that *FLO11* is highly polymorphic in length, varying from 3.0 to 6.1 kb. As expected, this length polymorphism is due to the expansion or contraction of a specific region of the *FLO11* gene that contains a 36 bp tandemly repeated consensus sequence (AACCCTKWCWACCCTGCAACCACYKCW-ACYACYGC). Thirteen different alleles containing 11, 28, 33, 44, 53, 56, 58, 61, 64, 69, 72, 75 and 78 repeats were seen (Fig. 4). The most frequent allele in the population contained 58 repeats, and this was seen in seven strains. Among the 20 strains tested, only six were heterozygous at

Table 3. *FLO11* polymorphism and *FLO11*-associated phenotypes

Strain	Biofilm mass (mg)	Mat diam. (cm)	Adhesion values (A_{570})	Hydrophobicity (%)	<i>FLO11</i> gene length (kb)	<i>FLO11</i> tandem repeats (kb)	<i>FLO11</i> tandem repeats* (kb)	No. of <i>FLO11</i> tandem repeats	<i>FLO11</i> tandem repeats melting temperature (°C)	<i>FLO11</i> relative expression
A68	0.320	7.53	1.83	97.39	5.6	3.3	—	78	81.6	10.25
V75	0.306	6.47	1.92	98.96	5.6	3.3	—	78	81.6	25.32
A28	0.313	6.63	1.91	97.58	5.6	3.2	—	75	81.5	14.06
A41	0.294	6.53	1.77	97.77	6.1	3.2	—	75	81.5	13.00
V80	0.182	5.00	1.38	88.55	6.1	3.1	—	72	81.4	0.25
V19	0.260	6.60	1.42	94.58	4.8	2.7	—	61	81.4	10.25
A43	0.276	7.07	2.04	96.54	5.0	2.6	—	58	81.3	26.01
A51	0.293	7.10	2.04	96.95	5.0	2.6	—	58	81.3	17.52
A9	0.278	7.73	2.02	93.35	5.0	2.6	—	58	81.3	19.04
V30	0.293	6.23	1.32	91.97	5.0	2.6	—	58	81.3	22.75
V44	0.285	5.83	1.35	93.37	5.0	2.6	—	58	81.3	6.95
V23	0.120	4.73	1.79	61.26	3.5	1.5	—	28	81.1	2.09
M12	0.221	4.43	2.44	98.22	3.1	0.9	—	11	81.0	2.39
M23	0.246	4.73	2.15	94.83	3.1	0.9	—	11	85.4	7.22
V16*	0.245	4.60	2.06	94.31	5.0	2.8	3.3	64	78	24.21
M39*	0.202	5.90	1.99	96.55	4.8	2.6	3.1	58	72	17.59
M66*	0.247	6.53	1.53	84.21	5.0	2.6	3.0	58	69	3.25
M46*	0.235	5.77	2.24	96.16	4.2	2.1	2.5	44	56	10.12
M49*	0.253	4.80	2.44	95.21	3.0	1.7	2.4	33	53	3.66
M25*	0.279	7.57	1.04	96.28	3.0	1.7	2.4	33	53	21.12
FM23†	0.420	4.42	0.21	39.04	3.1	0.9	—	11	85.4	59.71
FA9†	0.496	4.40	0.50	39.01	5.0	2.6	—	58	81.3	55.90

*Wild *S. cerevisiae* flor strains with two alleles at the *FLO11* locus.

†Flor strains derived from the laboratory strain BY4742 (this study).

Fig. 1. Mat formed on soft agar by A9, V80, V23 and M23 flor strains.

Fig. 2. Biofilm formed on flor medium by A9, V80, V23 and M23 flor strains.

the *FLO11* locus. The results of the high resolution melting analysis performed on strains homozygous at the *FLO11* locus showed that, except for strain M23, alleles of the same length have the same melting temperature and that the melting temperature generally increases with the number of repeats (Table 3).

The length of the *FLO11* repeated sequences correlates with biofilm-forming ability

It has been reported that in the same genetic background, differences in the lengths of repeated sequences in the *FLO11* gene correlate with variations in the intensities of some *FLO11*-associated phenotypes (Fidalgo *et al.*, 2008). To confirm this, we cloned the *FLO11* alleles from strains M23 and A9 into the BY4742 laboratory strain and analysed the *FLO11*-associated phenotypes of the FM23 and FA9 transformants obtained, respectively. For adhesion on plastic surfaces, FA9 showed adhesion of $0.50 \pm 0.1 A_{570}$ units, and FM23 0.21 ± 0.08 units. The adherence shown by the wild-type BY4742 strain was $0.20 \pm 0.02 A_{570}$ units, which was almost the same as that of FM23. For hydrophobicity, that of the wild-type BY4742 strain ($18.5 \pm 0.92\%$) was lower than those of FA9 ($39 \pm 0.9\%$) and FM23 ($39 \pm 0.3\%$). Also, BY4742 forms a small and smooth mat, while FM23 and FA9 exhibited larger mats with macroscopic structures. Finally, FA9 produced a biofilm of 0.496 mg and FM23 of

0.420 mg, whereas BY4742 did not form any detectable biofilm.

The same analysis regarding *FLO11* alleles and *FLO11*-dependent phenotypes was carried out considering different genetic backgrounds. To this end, Spearman's correlation coefficients (ρ) were calculated by considering the length of *FLO11* repeated sequences and mat formation, adherence to plastic, biofilm-forming ability and hydrophobicity of the 14 flor wild strains that were homozygous at the *FLO11* locus. For mat formation, the calculated ρ was 0.423 ($P=0.063$). For adherence to plastic surfaces, the calculated ρ was -0.415 ($P=0.071$). The value of ρ for the correlation of *FLO11* repeated sequences versus hydrophobicity was 0.312 ($P=0.184$). In all of these cases $P>0.05$, indicating no significant linear correlation between the phenotypes tested and the length of the *FLO11* repeated sequences at the 95.0% confidence level.

In contrast, the value of ρ calculated considering biofilm-forming abilities of these strains was 0.713 ($P=0.004$). This indicated a strong and significant linear correlation between biofilm-forming ability and length of the *FLO11* repeated sequences at the 95.0% confidence level. This linear correlation was also assessed by displaying the data as a scatter plot (Fig. 5a). Giving this strong correlation, the possibility of predicting the weight of the biofilm that would be formed (response variable) on the basis of the length of the *FLO11* repeats (predictor variable) was

Fig. 3. SEM images of strains A9, M23 and V80. (a) Biofilm cells embedded in extracellular matrix; (b) cells on the bottom of the flask.

verified by a simple regression model. The linear model that best fits the data is: biofilm formation (mg) = $0.16765 + 0.03812 \times FLO11$ repeat length (kb). In this model, the multiple R^2 value was 0.317 ($P=0.036$). In this model not all of the validity assumptions were satisfied, thus rendering the simple linear regression model proposed inappropriate.

***FLO11* transcription level and promoter sequence explain biofilm-forming ability**

Considering that biofilm formation may be influenced not only by the length of *FLO11* repeats, but also by the expression levels of the gene, we analysed the transcription level of *FLO11* and verified the possible correlation with

the weight of the biofilm formed in the 14 flor wild strains homozygous at the *FLO11* locus. While the value of ρ (0.581; $P=0.02$) indicated a significant correlation among *FLO11* transcription levels and the weights of the biofilm formed, the scatter plot revealed a non-linear relationship among the variables (Fig. 5b). In particular, small variations in *FLO11* transcription levels have a higher impact on the weight of the biofilm formed at the lower end of the scale, while variations above a certain value (around 10) do not have any effect on the phenotype. This suggests that a log-transformation of *FLO11* transcription values is appropriate to estimate a multiple regression linear model. The multiple regression model that best fits the data is: biofilm formation (mg) = $0.088236 + 0.022397 \times FLO11$ repeat length (kb) + $0.031168 \times \log(FLO11$ transcription level).

Fig. 4. *FLO11* repeated sequences of 20 wild flor strains of *S. cerevisiae*. Lanes: 1, 15, 16, 24, 1 kb DNA ladder (Sigma); 2–14, 17–23: *FLO11* repeated sequences of strains M12, M23, V23, M46, M66, A9, A43, A51, M39, V16, V19, V30, V44, A68, V80, V75, A28, A41, M25 and M49, respectively.

This model is highly significant as the multiple R^2 value was 0.751 ($P=0.0004$).

As the tested strains showed important differences in the induction of *FLO11* in flor medium, we analysed *FLO11* promoter sequences of strains showing high (A9), medium (M23), low (V23) and almost absent (V80) levels of *FLO11* transcription in flor medium. The lengths of the promoter regions sequenced were 2845 bp in V80, 2708 bp in M23, 2814 bp in V23 and 2738 in A9. The sequences of these promoters were subjected to multiple alignments with that of the *FLO11* gene of S288C, which showed substitutions from 3 to 5 % of the nucleotides. The four promoters were analysed in particular for the presence/absence of a 111 bp deletion, described by Fidalgo *et al.* (2006), localized in the upstream repression sequence at –1400 to –1200 upstream of ATG (Rupp *et al.*, 1999; Braus *et al.*, 2003).

The comparison across A9, M23, V23, V80 and S288C showed this deletion in A9, M23 and V23, according to Fidalgo *et al.* (2006), but not in V80 and S288C. In particular, the deletion occurred between nucleotides 1695 and 1799 (104 bp) in A9, 1695 and 1832 (137 bp) in M23, and 1678 and 1816 (138 bp) in V23. Other differences were also observed: A9 and V80 showed two repetitions of the short CAAATTAA sequence described by Fidalgo *et al.* (2006). Only one of these short sequences was seen in M23 and none in V23.

DISCUSSION

In this study, we have shown that in a natural population of 20 flor strains of *S. cerevisiae*, the *FLO11* gene is highly polymorphic, and that its allelic variability is mainly due to

Fig. 5. (a) Scatterplots of biofilm formed (weight) versus (a) length of *FLO11* repeats (kb) and (b) transcription levels of *FLO11* in 14 *S. cerevisiae* wild flor strains. *FLO11* transcription values in (b) are relative to transcription levels of the housekeeping gene *ACT1*.

the expansion and contraction of a central region containing different tandem repeats of a 36 bp consensus sequence. One interesting question is whether the generation and maintenance of these different alleles of the same gene offer wine yeasts an adaptive advantage in response to a particular environment, such as for must fermentation or biological ageing of wines.

We have further shown that in two transformants carrying two *FLO11* alleles that contain 11 and 58 repeats, the transcription levels were not significantly different. In these transformants, as the gene length increased, so did the weight of the biofilm. The same results were obtained by Fidalgo *et al.* (2008), who showed the phenotypic effects of different alleles of *FLO11* in flor/laboratory hybrids of *S. cerevisiae*. However, to conclude that the expansion and contraction of repeated sequences allow natural populations to rapidly adapt to a fluctuating environment, it is necessary to study the phenotypic effects of the different *FLO11* alleles in their specific genetic backgrounds. Accordingly, we studied different *FLO11*-associated phenotypes in 20 *S. cerevisiae* flor strains carrying *FLO11* alleles of different lengths. The strong and significant correlation seen only among the lengths of *FLO11* alleles and the biofilm-forming abilities of the strains homozygous at the *FLO11* locus, suggests that the length of *FLO11* has a greater role in biofilm formation on liquid surfaces than on other *FLO11*-associated phenotypes in flor strains, as hypothesized also by Fidalgo *et al.* (2008). However, the low multiple R^2 value of the simple regression model tested, suggests that other factors should be included in a model able to predict biofilm-forming ability. We hypothesized that one of these factors is the transcription level of *FLO11*.

The multiple linear regression model that includes transcription level and the size of repeats of *FLO11*, is highly significant. Thus, the air-liquid biofilm-forming ability of flor strains is dependent not only on the number of repeated sequences but also on the transcriptional levels of *FLO11*.

In order to understand the reasons behind the differences in transcription level among the flor strains, we analysed the promoter sequences of the *FLO11* alleles of strains A9, M23, V23 and V80. In the promoter of strain V80, the presence of the 111 bp sequence, localized in a well-characterized upstream repression sequence (Rupp *et al.*, 1999; Braus *et al.*, 2003), would explain why its *FLO11* transcription is repressed in flor medium and why strain V80 formed very little biofilm. On the contrary, this deletion is present in the promoter of *FLO11* in strains A9, M23 and V23, according to Fidalgo *et al.* (2006). Among the promoters analysed, we found two repeats of the short CAAATTAA sequence in A9 and V80. This sequence has been proposed for DNA intramolecular recombination and to be responsible for the evolutionary origin of the 111 bp deletion (Fidalgo *et al.*, 2006). The presence of only one copy of this sequence in strain M23 and none in V23

suggests that two direct repeats of the nucleotide sequence CAAATTAA are not sufficient to explain the evolutionary origin of this 111 bp deletion in *S. cerevisiae* flor strains.

Some studies have shown that *FLO11* expression is under epigenetic control as well as under genetic regulation (Verstrepen & Klis, 2006). Our results are consistent with the possibility that in the flor strains we have analysed here, epigenetic regulation has an important role in *FLO11* transcription during biofilm formation. Indeed, the differences in *FLO11* transcription levels among strains M23, V23 and A9 cannot be related solely to differences in their promoter sequences.

To our knowledge, this is the first study that proposes a multiple regression model that allows prediction of air-liquid biofilm-forming ability on the basis of transcription level and the length of the *FLO11* allele in a natural population of *S. cerevisiae* flor strains. Our results are consistent with the hypothesis that the ability to form a biofilm is as complex as other cell-cell adhesion phenotypes and might implicate a complete rearrangement of chromosomal DNA (Budroni *et al.*, 2005; Fidalgo *et al.*, 2006).

Finally, we documented the presence of an extracellular matrix in which cells from an air-liquid biofilm are embedded. An extracellular matrix has been observed in *S. cerevisiae* fluffy colonies by Kuthan *et al.* (2003), and recently by Beauvais *et al.* (2009) who reported that flocculating cells of *S. cerevisiae* secrete an extracellular matrix consisting of a mixture of glucose and mannose polysaccharides. The composition of the extracellular matrix associated with the air-liquid biofilm, and the possible correlation between *FLO11* expression and the amount of extracellular matrix, will be an intriguing puzzle for further investigations.

REFERENCES

- Barrales, R. R., Jimenez, J. & Ibeas, J. I. (2008). Identification of novel activation mechanisms for *FLO11* regulation in *Saccharomyces cerevisiae*. *Genetics* **178**, 145–156.
- Beauvais, A., Loussert, C., Prevost, M. C., Verstrepen, K. & Latgé, J. P. (2009). Characterization of a biofilm-like extracellular matrix in *FLO1*-expressing *Saccharomyces cerevisiae* cells. *FEMS Yeast Res* **9**, 411–419.
- Bowen, S. & Wheals, A. E. (2006). Ser/Thr-rich domains are associated with genetic variation and morphogenesis in *Saccharomyces cerevisiae*. *Yeast* **23**, 633–640.
- Bowen, S., Roberts, C. & Wheals, A. E. (2005). Patterns of polymorphism and divergence in stress-related yeast proteins. *Yeast* **22**, 659–668.
- Brachmann, C. B., Davies, A., Cost, G. J., Caputo, E., Li, J., Hieter, P. & Boeke, J. D. (1998). 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.
- Braus, G. H., Grundmann, O., Bruckner, S. & Mosch, H. U. (2003). Amino acid starvation and Gcn4p regulate adhesive growth and

- FLO11* gene expression in *Saccharomyces cerevisiae*. *Mol Biol Cell* **14**, 4272–4284.
- Budroni, M., Zara, S., Zara, G., Pirino, G. & Mannazzu, I. (2005). Peculiarities of flor strains adapted to Sardinian sherry-like wine ageing conditions. *FEMS Yeast Res* **5**, 951–958.
- Burke, D. A., Dawson, D. & Stearns, T. (2000). *Methods in Yeast Genetics: a Laboratory Manual*. Cold Spring Harbor, NY: Cold Spring Harbor Laboratory.
- Douglas, L. M., Li, L., Yang, Y. & Dranginis, A. M. (2007). Expression and characterization of the flocculin Flo11/Muc1, a *Saccharomyces cerevisiae* mannoprotein with homotypic properties of adhesion. *Eukaryot Cell* **6**, 2214–2221.
- Farris, G. A., Zara, S., Pinna, G. & Budroni, M. (2002). Genetic aspect of flor yeasts Sardinian strains, a case of study. In *Biodiversity and Biotechnology of Wine Yeasts*, pp. 71–83. Edited by Maurizio Ciani. Kerala, India: Research Signpost.
- Fidalgo, M., Barrales, R. R., Ibeas, J. I. & Jimenez, J. (2006). Adaptive evolution by mutations in the *FLO11* gene. *Proc Natl Acad Sci U S A* **103**, 11228–11233.
- Fidalgo, M., Barrales, R. R. & Jimenez, J. (2008). Coding repeat instability in the *FLO11* gene of *Saccharomyces* yeasts. *Yeast* **25**, 879–889.
- Frieman, M. B. & Cormack, B. P. (2004). Multiple sequence signals determine the distribution of glycosylphosphatidylinositol proteins between the plasma membrane and cell wall in *Saccharomyces cerevisiae*. *Microbiology* **150**, 3105–3114.
- Halme, A., Bumgarner, S., Styles, C. A. & Fink, G. R. (2004). Genetic and epigenetic regulation of the *FLO* gene family generates cell-surface variation in yeast. *Cell* **116**, 405–415.
- Ishigami, M., Nakagawa, Y., Hayakawa, M. & Iimura, Y. (2004). *FLO11* is essential for flor formation caused by the C-terminal deletion of *NRG1* in *Saccharomyces cerevisiae*. *FEMS Microbiol Lett* **237**, 425–430.
- Kuchin, S., Vyas, V. K. & Carlson, M. (2002). Snf1 protein kinase and the repressors Nrg1 and Nrg2 regulate *FLO11*, haploid invasive growth, and diploid pseudohyphal differentiation. *Mol Cell Biol* **22**, 3994–4000.
- Kuthan, M., Devaux, F., Janderová, B., Slaninová, I., Jacq, C. & Palková, Z. (2003). Domestication of wild *Saccharomyces cerevisiae* is accompanied by changes in gene expression and colony morphology. *Mol Microbiol* **47**, 745–754.
- Liu, H., Styles, C. A. & Fink, G. R. (1996). *Saccharomyces cerevisiae* S288C has a 20 mutation in *FLO8*, a gene required for filamentous growth. *Genetics* **144**, 967–978.
- Lo, W. S. & Dranginis, A. M. (1998). The cell surface flocculin Flo11 is required for pseudohyphae formation and invasion by *Saccharomyces cerevisiae*. *Mol Biol Cell* **9**, 161–171.
- Mannazzu, I., Simonetti, E., Marinangeli, P., Guerra, E., Budroni, M., Thangavelu, M., Bowen, S., Wheals, A. & Clementi, F. (2002). *SED1* gene length and sequence polymorphisms in feral strains of *Saccharomyces cerevisiae*. *Appl Environ Microbiol* **68**, 5437–5444.
- Marinangeli, P., Clementi, F., Ciani, M. & Mannazzu, I. (2004). *SED1* polymorphism within the genus *Saccharomyces*. *FEMS Yeast Res* **5**, 73–79.
- Pan, X. & Heitman, J. (1999). Cyclic AMP-dependent protein kinase regulates pseudohyphal differentiation in *Saccharomyces cerevisiae*. *Mol Cell Biol* **19**, 4874–4887.
- Pan, X. & Heitman, J. (2002). Protein kinase A operates a molecular switch that governs yeast pseudohyphal differentiation. *Mol Cell Biol* **22**, 3981–3993.
- Pfaffl, M. W. (2001). A new mathematical model for relative quantification in real-time RT-PCR. *Nucleic Acids Res* **29**, e45.
- R Development Core Team (2008). R: A language and environment for statistical computing. Vienna: R Foundation for Statistical Computing. www.R-project.org.
- Reynolds, T. B. & Fink, G. R. (2001). Bakers yeast, a model for fungal biofilm formation. *Science* **291**, 878–881.
- Rice, P., Longden, I. & Bleasby, A. (2000). EMBOS: the European molecular biology open software suite. *Trends Genet* **16**, 276–277.
- Rupp, S., Summer, S. E., Lo, H. J., Madhani, H. & Fink, G. R. (1999). MAP kinase and cAMP filamentation signaling pathways converge on the unusually large promoter of the yeast *FLO11* gene. *EMBO J* **18**, 1257–1269.
- Sherman, F., Fink, G. & Hicks, J. (1986). *Methods in Yeast Genetics*. Cold Spring Harbor, NY: Cold Spring Harbor Laboratory.
- Stephens, A. J., Inman-Bamber, J., Giffard, P. M. & Huygens, F. (2008). High-resolution melting analysis of the *spa* repeat region of *Staphylococcus aureus*. *Clin Chem* **54**, 432–436.
- Verstrepen, K. J. & Klis, F. M. (2006). Flocculation, adhesion and biofilm formation in yeasts. *Mol Microbiol* **60**, 5–15.
- Verstrepen, K. J., Jansen, A., Lewitter, F. & Fink, G. R. (2005). Intragenic tandem repeats generate functional variability. *Nat Genet* **37**, 986–990.
- Zara, S., Bakalinsky, A. T., Zara, G., Pirino, G., Demontis, M. A. & Budroni, M. (2005). *FLO11*-based model for air-liquid interfacial biofilm formation by *Saccharomyces cerevisiae*. *Appl Environ Microbiol* **71**, 2934–2939.

Edited by: M. Schweizer