Creation of stable heterothallic strains of *Komagataella phaffii* enables dissection of mating gene regulation

Lina Heistinger, Brigitte Gasser, Diethard Mattanovich

Supplemental material

FIG S1 Mating plates.









FIG S1 Mating plates. To analyse the effect of different mutations, strains were crossed on mating agar and replicated onto selective agar plates to evaluate diploid cell formation. Representative images of all tested combinations are shown. To cross two strains with the same mutation, additional marker genes were integrated for selection. (A) Mating of homothallic *mat* Δ strains with each other and the wildtype. Please note that the two wildtype strains on each plate carry different antibiotic markers for the selection of diploid clones with both mutant strains present on the same plate (1-ZeoR, 2-hphMX, 3-kanMX, 4-NatMX). (B) Mating of *ste2* Δ and *ste3* Δ strains. (C) Mating of heterothallic *dic1-2* Δ mat Δ strains. (D) Mating of heterothallic *dic1-2* Δ strains with the unmodified *dic1-2* Δ strains.

strains gene	$dic1-2\Delta(\alpha)/dic1-2\Delta(\mathbf{a})$	dic1-2 $\Delta(\alpha)$ /wt	<i>dic1-2</i> ∆(a)/wt
MATa1	0.0018	0.0235	0.1377
MATa2	0.0061	0.0212	0.0056
MATα1	0.0057	0.1714	0.0047
ΜΑΤα2	0.0001	0.0935	0.0055
STE2	0.0036	0.0117	0.0003
STE3	0.0037	0.0043	0.0019

 Table S1 p-values for comparison of expression levels in mating medium

All p-values were calculated from three biological replicates using a Student's t-test.

Table S2 p-values for comparison of expression levels in YPD vs. mating medium

strain gene	dic1-2 $\Delta(\alpha)$	<i>dic1-2</i> ∆(a)	wildtype
MATa1	0.0356	0.0127	0.0460
MATa2	-	0.0064	0.0209
ΜΑΤα1	0.0058	-	0.0047
ΜΑΤα2	0.0001	-	0.0058
STE2	-	0.0036	0.0109
STE3	0.0033	0.0074	0.0018

All p-values were calculated from three biological replicates using a Student's t-test. Values are not shown for comparisons in which transcript levels were extremely low in both conditions.

		MATa1			MATa2			MATa1			ΜΑΤα2			STE2			STE3	
strain	fold change	SD	p-value															
<i>dic1-2</i> ∆(a)	1.0	+/- 0.10	-	1.0	+/- 0.16	-	1.0	+/- 0.08	-	1.0	+/- 0.31	-	1.0	+/- 0.16	-	1.0	+/- 0.19	-
<i>mat</i> a1∆(a)	-	-	-	1.16	+/- 0.15	0.1282	0.95	+/- 0.23	0.7183	0.66	+/- 0.17	0.0494	1.14	+/- 0.44	0.7589	1.20	+/- 0.26	0.2013
mat a 2∆(a)	1.76	+/- 0.05	0.0000	-	-	-	0.60	+/- 0.08	0.0001	0.67	+/- 0.41	0.1669	0.01	+/- 0.00	0.0000	0.19	+/- 0.03	0.0000
$mat \alpha 1 \Delta(\mathbf{a})$	1.05	+/- 0.28	1.0000	1.27	+/- 0.33	0.2324	-	-	-	6.01	+/- 0.48	0.0000	1.91	+/- 0.50	0.0049	1.12	+/- 0.25	0.4547
$mat\alpha 2\Delta(\mathbf{a})$	1.35	+/- 0.18	0.0022	1.10	+/- 0.29	0.6250	1.08	+/- 0.04	0.0022	-	-	-	1.40	+/- 0.23	0.0056	2.02	+/- 0.20	0.0000
<i>dic1-2</i> Δ(α)	1.0	+/- 0.14	-	1.0	+/- 0.17	0.0596	1.0	+/- 0.21	-	1.0	+/- 0.20	-	1.0	+/- 0.13	-	1.0	+/- 0.22	-
$mata1\Delta(\alpha)$	-	-	-	0.78	+/- 0.20	-	1.10	+/- 0.06	0.3423	1.25	+/- 0.15	0.0486	1.12	+/- 0.16	0.2168	1.11	+/- 0.07	0.3015
mata $2\Delta(\alpha)$	1.37	+/- 0.07	0.0045	-	-	-	1.04	+/- 0.08	0.7193	1.12	+/- 0.07	0.2217	0.89	+/- 0.14	0.2354	0.98	+/- 0.10	0.8228
$mat \alpha 1 \Delta(\alpha)$	0.75	+/- 0.27	0.0724	0.44	+/- 0.20	0.0045	-	-	-	1.44	+/- 0.37	0.0339	1.03	+/- 0.21	0.9094	0.01	+/- 0.00	0.0000
$mat\alpha 2\Delta(\alpha)$	0.86	+/- 0.16	0.1329	0.94	+/- 0.26	0.5183	0.46	+/- 0.08	0.0000	-	-	-	1.10	+/- 0.06	0.1536	1.00	+/- 0.06	0.9808

Table S3 Relative *MAT* and *STE2 / STE3* transcript levels in *dic1-2* Δ *mat* Δ strains.

Changes in *MAT* and *STE2 / STE3* transcript levels in the different $dic1-2\Delta mat\Delta$ strains under mating conditions were analyzed by quantitative PCR. Gene expression was normalized to *ACT1* and is given as fold change to the expression in the respective unmodified $dic1-2\Delta$ strain. Data was analyzed using the $\Delta\Delta C_{T}$ method and fold changes were calculated as $2^{-\Delta\Delta CT}$. Fold change values of highly expressed genes transcribed from the active *MAT* locus are indicated in bold. P-values were calculated using a Student's t-test.



1. PCR amplification of fragmements introducing overlapping ends 1. PCR amplification of fragmements with fusion sites for Golden Gate Cloning 1. PCR amplification of fragmements with fusion sites for Golden Gate Cloning



FIG S2 Schematic of different strategies for the generation of the split-marker knockout cassettes. Generally, the knockout cassettes consisted of the homologous region for genomic integration (fragment A + D shown in green) and two halves of an antibiotic resistance marker with overlapping ends of around 440 bp (B + C shown in magenta). In the first approach, the separate fragments were amplified by PCR introducing small overlapping sequences and the final cassettes were generated by a second fusion PCR step. For the two remaining strategies, the fragments were introduced into plasmid vectors by Golden Gate Cloning. PCR amplified fragments carrying the required fusion sites were either introduced into separate vectors for the two halves of the split marker cassette or cloned as the full knockout cassette containing the whole marker sequence. In both cases the final split-marker fragments were amplified by PCR using the purified vectors as templates.

Table S4 List of primers used in this study.

knockout cassettes assembl	ed by fusion PCR of A-B and C-D	
MATal	homologous region A	AACTTGCTCTCCGATTTG
		GTTGTCGACCTGCAGCGTACCTTGGGTCAGATGTGCGT
	marker fragment B	ACGCACATCTGACCCAAGGTACGCTGCAGGTCGACAAC
		CTGCTCGCCGATCTCGGTC
	marker fragment C	TGACCAGTGCCGTTCCGG
		TTACGTTATGGATGGGACTAGTGGATCTGATATCACCTA
	homologous region D	TAGGTGATATCAGATCCACTAGTCCCATCCATAACGTAAC
		CCGTGATGTAGCCGAAGAG
MATa1+2	homologous region A	GTGATGTAGCCGAAGAGT
		GTTGTCGACCTGCAGCGTACATCCTGGAATGAGCTACT
	marker fragment B	AGTAGCTCATTCCAGGATGTACGCTGCAGGTCGACAAC
		CGGTGAGAATGGCAAAAGCTTATG
	marker fragment C	AAGCCCGATGCGCCAGAGTTG
		CAGTTTTCATGTCTCAAGCGAGTGGATCTGATATCACCTA
	homologous region D	TAGGTGATATCAGATCCACTCGCTTGAGACATGAAAACTG
		CCAGCCAAACAAACACAATC
knockout cassettes assembl	led into 2 different vectors by Golden Gate Clon	ing (A-B and C-D)
MATa2	homologous region A (fusion sites A-B)	GGTCTCCGATCTTTTGGTATGGAGGTGCCG
		GGTCTCCCCGGACAACGTTTAGCTGCCTTC
	marker fragment B (fusion sites B-C)	GGTCTCC CCGG AGTGGATCTGATATCACCTA
		GGTCTCCAATTAGACCTGCCTGAAACCGAACTGC
	marker fragment C (fusion sites A-B)	GGTCTCCGATCAGATGTTGGCGACCTCGTATTG
		GGTCTCCCCGGGTACGCTGCAGGTCGACAAC
	homologous region D (fusion sites B-C)	GGTCTCCCCGGAATTCTATCCATCAGACCTG
		GGTCTCCAATTTGCCTGTTGCTTCTATACT
MATal	homologous region A (fusion sites A-B)	GGTCTCCGATCGCTCAGCGACCAAGAAAA
		GGTCTCCCCGGAGCTACTCAGCACTAGAAA
	marker fragment B (fusion sites B-C)	GGTCTCCCCGGGTACGCTGCAGGTCGACAAC
		GGTCTCCAATTCGGTGAGAATGGCAAAAGCTTATG
	marker fragment C (fusion sites A-B)	GATAGGTCTCCGATCAAGCCCGATGCGCCAGAGTTG
		GATAGGTCTCCCCGGAGTGGATCTGATATCACCTA
	homologous region D (fusion sites B-C)	GATAGGTCTCCCCGGTATCTATCACATGCACGCAC
		GATAGGTCTCCAATTTTCCATCCTTTATTCGCCT
ΜΑΤα2	homologous region A (fusion sites A-B)	GGTCTCCGATCCTCTACAGAATGCGCGAA
		GGTCTCCCCGGACAGAGCTACTTCAAGCA
	marker fragment B (fusion sites B-C)	GGTCTCCCCGGGTACGCTGCAGGTCGACAAC
		GGTCTCCAATTTGCGTTGACGTTGGTGAC
	marker fragment C (fusion sites A-B)	GGTCTCCGATCGTCCTTCACCACCGACACC
		GGTCTCC CCGG AGTGGATCTGATATCACCTA
	homologous region D (fusion sites B-C)	GGTCTCC CCGG AACTGATAAGAATCGCGG
		GGTCTCCAATTCCAGCCAAACAAACAAAA
knockout cassettes assembl	led into 1 vector by Golden Gate Cloning (A-D)	1
MATal+2	homologous region A (fusion sites A-B)	GATAGGTCTCCGATCTTTTGGTATGGAGGTGCCG
		GATAGGTCTCCCCGGACAACGTTTAGCTGCCTTC

	marker fragment B-C (fusion sites B-C)	GATAGGTCTCCCCGGGTACGCTGCAGGTCGACAAC
		GATAGGTCTCCAATTAGTGGATCTGATATCACCTA
	homologous region D (fusion sites C-D)	GATAGGTCTCCAATTAGTCCCATCCATAACGTAAG
		GATAGGTCTCCAGCTCCGTGATGTAGCCGAAGAG
STE2	homologous region A (fusion sites A-B)	GATAGGTCTCCGATCAAATCTTCAGCAAGAACGG
		GATAGGTCTCCCCGGGCAGGCAATAATCAAACGG
	marker fragment B-C (fusion sites B-C)	GATAGGTCTCCCCGGGTACGCTGCAGGTCGACAAC
		GATAGGTCTCCAATTAGTGGATCTGATATCACCTA
	homologous region D (fusion sites C-D)	GATAGGTCTCCAATTGCACGTTCACTTCTCTTTTACT
		GATAGGTCTCCAGCTACTGAAACACCAAAAGATCCC
STE3	homologous region A (fusion sites A-B)	GATAGGTCTCCGATCGGAAGCCAAGGGGAATGA
		GATAGGTCTCCCCGGGATAAAGTGTGACGCGGA
	marker fragment B-C (fusion sites B-C)	GATAGGTCTCCCCGGGTACGCTGCAGGTCGACAAC
		GATAGGTCTCCAATTAGTGGATCTGATATCACCTA
	homologous region D (fusion sites C-D)	GATAGGTCTCCAATTAACAGGCCAAACATGAAGA
		GATAGGTCTCCAGCTCTCGCATTAACCATTTTCTCC
DIC1-2	homologous region A (fusion sites A-B)	GATAGGTCTCCGATCTGAAAGATTTGGGTTGCG
		GATAGGTCTCCCCGGGATGATGTTCCTGGTTTACT
	marker fragment B-C (fusion sites B-C)	GATAGGTCTCCCCGGGTACGCTGCAGGTCGACAAC
		GATAGGTCTCCAATTAGTGGATCTGATATCACCTA
	homologous region D (fusion sites C-D)	GATAGGTCTCCAATTCATCATCGTTTTCCCACAAG
		GATAGGTCTCCGAGACTCTGGAGGTATATG
		GATAGGTCTCCTCTCAAGGAGCAATCATACC
		GATAGGTCTCCAGCTTCCCAAATACGCACACAG
MATa1+2 integration	homologous region A (fusion sites A-B)	GATAGGTCTCCGATCTAACGGGGGATCATGTATCGG
		GATAGGTCTCCCCGGTTATTCATCCGCATGGTATCCA
	marker fragment B-C (fusion sites B-F)	GATAGGTCTCCCCGGGTACGCTGCAGGTCGACAAC
		GATAGGTCTCCAAGCAGTGGATCTGATATCACCTA
	<i>MATa1+2</i> (fusion sites F-C)	GATAGGTCTCCGCTTGTAACTGAGTTGAATAGCTTACG
		GATAGGTCTCCAATTCATTTAGTATTTGACGCTACAGG
	homologous region D (fusion sites C-D)	GATAGGTCTCCAATTCATCATCGTTTTCCCACAAG
		GATAGGTCTCCGAGACTCTGGAGGTATATG
		GATAGGTCTCCTCTAAGGAGCAATCATACC
		GATAGGTCTCCAGCTTCCCAAATACGCACACAG
<i>MATa1+2</i> integration	homologous region A and $MAT\alpha 1+2$ (fusion sites A-B)	GATAGGTCTCCGATCTTTTGGTATGGAGGTGCCG
		GATAGGTCTCCCCGGGCAGATTATTAATGAAGTATAGAAACC
	marker fragment B-C (fusion sites B-C)	GATAGGTCTCCCCGGGTACGCTGCAGGTCGACAAC
		GATAGGTCTCCAATTAGTGGATCTGATATCACCTA
	homologous region D (fusion sites C-D)	GATAGGTCTCCAATTTTACTCTTCGTCCGCATGGT
		GATAGGTCTCCAGCTTAACGGGGGATCATGTATCGG
primers for cloning of ho	mology templates for Cas9 mediated integratio	n
<i>MATal</i>	homologous region A (fusion sites A-B)	GATAGGTCTCCGATCAAGGCAGCTAAACGTTGT
		GATAGGTCTCCCCGGTGCCAAGACTTGTGTTTCT
	marker fragment B-C (fusion sites B-C)	GATAGGTCTCCCCGGGTACGCTGCAGGTCGACAAC
		GATAGGTCTCCAATTAGTGGATCTGATATCACCTA
	homologous region D (fusion sites C-D)	GATAGGTCTCCAATTTTACACAAAGCCAGCAAC

		GATAGGTCTCCAGCTATGGGACTGAGCTTGAAG
MATa2	homologous region A (fusion sites A-B)	GATAGGTCTCCGATCAAGGCAGCTAAACGTTGT
		GATAGGTCTCCCCGGTGCCAAGACTTGTGTTTCT
	marker fragment B-C (fusion sites B-C)	GATAGGTCTCCCCGGGTACGCTGCAGGTCGACAAC
		GATAGGTCTCCAATTAGTGGATCTGATATCACCTA
	homologous region D (fusion sites C-D)	GATAGGTCTCCAATTCTAACTAGTTATCATACTTCTGTCCTTTCTACC
		GATAGGTCTCCAGCTTGTTTTCCCGAATAGGCT
ΜΑΤαΙ	homologous region A (fusion sites A-B)	GATAGGTCTCCGATCATTTCTAGTGCTGAGTAGC
		GATAGGTCTCCCCGGCCTTTCAAGTGTTCCCAA
	marker fragment B-C (fusion sites B-C)	GATAGGTCTCCCCGGGTACGCTGCAGGTCGACAAC
		GATAGGTCTCCAATTAGTGGATCTGATATCACCTA
	homologous region D (fusion sites C-D)	GATAGGTCTCCAATTCTAACTAGTTAACTCCGATTGCCTTTTGT
		GATAGGTCTCCAGCTGAATGCTCAAGAACGACG
ΜΑΤα2	homologous region A (fusion sites A-B)	GATAGGTCTCCGATCACTCCGATTGCCTTTTG
		GATAGGTCTCCCCGGGAGAAAAAGAAGGATGCC
	marker fragment B-C (fusion sites B-C)	GATAGGTCTCCCCGGGTACGCTGCAGGTCGACAAC
		GATAGGTCTCCAATTAGTGGATCTGATATCACCTA
	homologous region D (fusion sites C-D)	GATAGGTCTCCAATTTTTCCAACATTAAAGCGATCG
		GATAGGTCTCCTAACTAGTTAGAGAATCTTGCAAGGGATATC
		GATAGGTCTCCGTTACATCAGAGTTTTGTGCACA
		GATAGGTCTCCAGCTGCTACAGAAAACCGCGA
primers for assembly of g	gRNAs by fusion PCR	
for all a DNA -	CD with 1 aDNA-11 area	CGCCATGCCGAAGCATGTTGCCCAGCCGGCGCCAGCGAGGAGGCTGG
jor au gRNAs	CR_rib_l_gRNAall_rev	GACCATGCCGGCC
	CR rib 2 gRNAall rev	GATAGGTCTCCAAGCGTCCAAAGCTGTCCCATTCGCCATGCCGAAGCA
		TGTTGCCCAGCC
	CR_rib_3_gRNAall_rev	AGGCTGGGACCATGCCGGCCAAAAGCACCGACTCGGTGCCACTTTTTC
		AAGTTGATAACG
	CR_rib_4_gRNAall_fwd	ACITITICAAGIIGATAACGGACIAGCCITATITITAACIIGCTATITCTA
MATal	MATa1_gRNA_1_fwd	GTAAGCTCGTCTA
		AAACGAGTAAGCTCGTCTATCATGGACAATAGGCATGGTTTTAGAGCT
	MATa1_gRNA_2_fwd	AGAAATAGCAAG
		GATAGGTCTCCCATGATTCTGCTGATGAGTCCGTGAGGACGAAACGAG
MATa2	MATa2_gRNA_1_fwd	TAAGCTCGTCCA
	MAT-2 -DNA 2 feed	AAACGAGTAAGCTCGTCCAGAATCAAAGGATGACGATGTTTTAGAGCT
	MATA2_gRNA_2_1wd	AGAAATAGCAAG
MATal	MATalpha1 gRNA1 fwd	GATAGGTCTCCCATGTTGCGGCTGATGAGTCCGTGAGGACGAAACGAG
1917 LI WI	hin ri uphur_pre h ri_i wa	TAAGCTCGTCCC
	MATalpha1_gRNA2_fwd	AAACGAGTAAGCTCGTC <u>CCGCAACCCAAATCTTTCAA</u> GTTTTAGAGCT
		AGAAATAGCAAG
ΜΑΤα2	MATalpha2_gRNA_1_fwd	GATAGGTCTCCCATGTTTCTCCTGATGAGTCCGTGAGGACGAAACGAG
	MATalpha2_gRNA_2_fwd	
nrimers for cassotte ama	lification and integration control	
MATal	MATal control fud	TGTAAGAGGAGAGTGGAATTG
1717 1 1 1		
		AUUUAUIUUIAAUAIIUU

MATa2	MATa2_ampl_fwd	TTTTGGTATGGAGGTGCCG
	HphMX_rev	AGATGTTGGCGACCTCGTATTG
	HphMX_fwd	AGACCTGCCTGAAACCGAACTGC
	MATa2_ampl_rev	TGCCTGTTGCTTCTATACT
	MATa2_control_fwd	CAATCCTCCATGACTCAGA
	MATa2_control_rev	ATAACGAGATGACTGACGG
MATal+2	MATa12_ampl_fwd	TTTTGGTATGGAGGTGCCG
	Zeo_rev	CTGCTCGCCGATCTCGGTC
	Zeo_fwd	TGACCAGTGCCGTTCCGG
	MATa12_ampl_rev	CCGTGATGTAGCCGAAGAG
-	MATa12_control_fwd = Zeo_fwd	TGACCAGTGCCGTTCCGG
-	MATa12_control_rev = MATa1_control_rev	AGGGAGTCGTAAGATTGG
ΜΑΤα1	MATa1_ampl_fwd	GCTCAGCGACCAAGAAAA
	KanMX_rev	CGGTGAGAATGGCAAAAGCTTATG
	KanMX_fwd	AAGCCCGATGCGCCAGAGTTG
	MATal_ampl_rev	TTCCATCCTTTATTCGCCT
	MATa1_control_fwd	GTGATGTAGCCGAAGAGT
	MATal_control_rev	AGAATGCTCAAGAACGAC
ΜΑΤα2	MATα2_ampl_fwd	CTCTACAGAATGCGCGAA
	NatMX_rev	TGCGTTGACGTTGGTGAC
	NatMX_fwd	GTCCTTCACCACCGACACC
	MATα2_ampl_rev	CCAGCCAAACAAACACAA
	MATa2_control_fwd	CTAGTGCGAGGAGTTTCT
	MATa2_control_rev	CTTCTTCTCGTTTCCCACC
MATal+2	MATa12_control_fwd	GTAAGGGAGTCGTAAGATTGG
	MATa12_control_rev	AGTGCAGGAGGAAACAAG
STE2	STE2_ampl_fwd	AAATCTTCAGCAAGAACGG
	KanMX_rev	CGGTGAGAATGGCAAAAGCTTATG
	KanMX_fwd	AAGCCCGATGCGCCAGAGTTG
	STE2_ampl_rev	ACTGAAACACCAAAAGATCCC
	STE2_control_fwd	CGACAGAAGGGAAGCAAG
	STE2_control_rev	AAAAGAGATAACGTCCAGCG
STE3	STE3_ampl_fwd	GGAAGCCAAGGGGAATGA
	NatMX_rev	TGCGTTGACGTTGGTGAC
	NatMX_twd	GICCTICACCACCGACACC
	STE3_ampl_rev	
	STE3_control_twd	
DIGLA	SIE3_control_rev	
DIC1-2	DIC1_ampl_fwd	
	KanMA_rev	
	KanMA_fwd	
	DIC1_ampl_rev	
	DIC1_control_twd	
MATULALA	DIC1_control_rev	
MATa1+2 integration	MA1a12_integ_ampl_fwd	
	NatMX_rev	
	NatMX_fwd	GTUUTTCACCACUGACACC

	MATa12_integ_ampl_fwd = DIC1_ampl_rev	TCCCAAATACGCACACAG
	MATa12_integ_control_fwd	AACGCAGGAGAACCGTAAAC
	MATa12_integ_control_rev =	GTGTCTCGAGTTTTGTGTAGT
	DIC1_control_rev	
MATa1+2 integration	MATa12_integ_ampl_fwd	TTTTGGTATGGAGGTGCCG
	NatMX_rev	TGCGTTGACGTTGGTGAC
	NatMX_fwd	GTCCTTCACCACCGACACC
	MATa12_integ_ampl_rev	TAACGGGGATCATGTATCGG
	MATa12_integ_control_fwd	GGAATAAAGCGTACTGGTG
	MATa12_integ_control_rev	GCTCTACAGAATGCGCGAA
primers for mating-type de	termination	
MAT locus1	MAT_locus1_preDIC1_fwd	GCATCCAGCATAATTTAGATT
	MAT_locus1_MATa_rev	AACTCTGGGATCTTTGGA
	MAT_locus1_MATa_rev	CGCTTGAGACATGAAAACTG

Fusion sites for Golden Gate Cloning are indicated in bold. Underlined characters indicate guide RNA

sequences for CRISPR/Cas9 mediated homology-directed repair.