

Supplemental material

Natural DNA transformation is functional in *Lactococcus lactis* ssp. *cremoris* KW2

Blandine David, Amandine Radziejwoski, Frédéric Toussaint, Laetitia Fontaine, Marie Henry de Frahan, Cédric Patout, Sabine van Dillen, Patrick Boyaval, Philippe Horvath, Christophe Fremaux, Pascal Hols

TABLE S1. Late *com* genes in the complete genomes of *L. lactis* ssp. *cremoris*.

TABLE S2. Up-regulated genes in the ComX⁺ strain (BLD101 [pGIBLD001]) compared to Ctl⁻ (BLD101 [pG⁺host9]).

TABLE S3. Predicted Com-box in intergenic regions of the genome of KW2.

FIGURE S1. Protein and nucleotide mutations of a selected spontaneous Str^r mutant of MG1363.

FIGURE S2. Stabilization of the activation of the late promoter P_{comGA} by inducible *comX* expression.

TABLE S1. Late *com* genes in the complete genomes of *L. lactis* ssp. *cremoris*.

Gene	KW2			MG1363			A76			SK11			UC509.9		
	Sta-tus ^a	accession number	size (aa)	Sta-tus ^a	accession number	size (aa)/id. (%) ^b	Sta-tus ^a	accession number	size (aa)/id. (%) ^b	Sta-tus ^a	accession number	size (aa)/id. (%) ^b	Sta-tus ^a	accession number	size (aa)/id. (%) ^b
<i>comX</i>	+	AGV74143	163	+	CAL98991	163/99	+	AEU41666	163/98	+	ABJ73891	116/96	+	AFW92540	163/99
<i>comGA</i>	+	AGV74123	326	+	CAL98972	326/100	Tn	AEU41642	180/99	Tn	NA ^e	295/99	+	AFW92525	326/99
<i>comGB</i>	+	AGV74122	314	+	CAL98971	362 ^c /100	+	AEU41641	310 ^d /98	*	NA		+	AFW92524	310 ^d /98
<i>comGC</i>	+	AGV74121	125	+	CAL98970	125/98	+	AEU41640	125/98	+	ABJ73870	125/98	+	AFW92523	125/98
<i>comGD</i>	+	AGV74120	132	+	CAL98969	147 ^c /98	*	AEU41639 & AEU41638	55/97 & 62/94	*	NA		*	NA	
<i>comGE</i>	+	AGV74119	98	+	CAL98968	128 ^c /100	+	AEU41637	78 ^d /96	+	ABJ73869	98/94	+	AFW92522	128 ^c /96
<i>comGF</i>	+	AGV74118	141	+	CAL98967	148 ^c /100	+	AEU41636	86 ^d /100	+	ABJ73868	141/99	+	AFW92521	148 ^c /99
<i>comGG</i>	+	AGV74117	99	+	CAL98966	99/98	+	AEU41635	81/98	+	ABJ73867	99/98	+	NA ^c	99/98
<i>comC</i>	+	AGV74021	221	+	CAL98868	221/99	+	AEU41516	221/99	*	ABJ73767	195/98	*	AFW92429	195/98
<i>comFA</i>	+	AGV72992	430	+	CAL98061	430/98	*	AEU40370 & AEU40371	163/99 & 261/97	*	NA		*	NA	
<i>comFC</i>	+	AGV72991	216	+	CAL98062	216/98	*	AEU40369	124/98	+	ABJ72723	216/98	+	AFW91637	216/97
<i>comEA</i>	+	AGV73706	218	+	CAL98523	204 ^d /98	*	AEU39881 & AEU39882	38/90 & 193/93	*	ABJ73427	200/91	*	AFW92184	200/89
<i>comEC</i>	+	AGV73705	736	*	NA		Tn	AEU39883 & AEU39887	395/97 & 75/97	Tn	NA	466/97	Tn	AEU39887	75/97
<i>ssbB</i>	+	AGV72374	129	+	CAL97013	129/99	+	AEU39634	129/97	+	ABJ72042	129/97	+	AFW91073	129/95
<i>coiA</i>	+	AGV73660	327	*	NA		+	AEU39940	327/97	+	ABJ73377	327/97	+	AFW92139	327/97
<i>dprA</i>	+	AGV73150	282	+	CAL97866	282/98	+	AEU40490	282/97	+	ABJ72831	282/98	+	AFW91767	282/98
<i>RecA</i>	+	AGV72340	387	+	CAL96979	387/100	+	AEU39591	387/100	+	ABJ72007	387/100	+	AFW91039	387/100

^a +, presence of a complete gene; * gene incomplete due to nucleotide(s) exchange, insertion or deletion; Tn, gene disrupted by the insertion of at least one transposon.

^b Percentage of identity (id.) with proteins of strain KW2.

^c Start codon annotated upstream.

^d Start codon annotated downstream.

^e NA, non-annotated, protein was manually identified.

TABLE S2. Up-regulated genes in the ComX⁺ strain (BLD101 [pGIBLD001]) compared to Ctl⁻ (BLD101 [pG⁺host9]).

Locus tag	Gene	Description	FC ^a	Essential late <i>com</i> ^b	Com-box ^c
<i>comX</i>					
kw2_2209	<i>comX</i>	competence regulator ComX	19.8	Y	
Com-box-associated late loci					
<i>Locus I</i>					
kw2_0357	<i>recA</i>	recombination protein A RecA	2.8	Y	Y
<i>Locus II</i>					
kw2_0391	<i>ssbB</i>	single-strand binding protein SsbB	23.9	Y	Y
kw2_0392	<i>groES</i>	chaperonin GroES	4.8		
kw2_0393	<i>groEL</i>	chaperonin GroEL	3.6		
<i>Locus III</i>					
kw2_1003	<i>radC</i>	DNA repair protein RadC	3.0		Y
kw2_1004	<i>rex</i>	redox-sensing transcriptional repressor Rex	2.8		
kw2_1005	<i>arsC</i>	ArsC family protein	2.5		
<i>Locus IV</i>					
kw2_1028		PhoH family protein	2.8		
kw2_1029		MutT/NUDIX family protein	1.6		
kw2_1031		diacylglycerol kinase	2.2		
kw2_1032		hypothetical protein	1.7		
kw2_1033	<i>comFC</i>	competence protein FC	1.7	Y	
kw2_1034	<i>comFA</i>	competence protein FA	1.9	Y	Y
<i>Locus V</i>					
kw2_1192	<i>dprA</i>	DNA protecting protein DprA	6.2	Y	Y
kw2_1193	<i>topA</i>	DNA topoisomerase I TopA	2.3		
<i>Locus VI</i>					
kw2_1576	<i>ffh</i>	signal recognition particle protein Ffh	3.7		
kw2_1577		cell surface protein	3.4		Y
<i>Locus VII</i>					
kw2_1701	<i>alaS</i>	alanyl-tRNA synthetase AlaS	1.6		
kw2_1702	<i>prsA</i>	peptidyl-prolyl cis-trans isomerase	2.5		
kw2_1703		O-methyltransferase family protein	3.4		
kw2_1704	<i>pepF</i>	oligoendopeptidase F PepF	3.8		
kw2_1705	<i>coiA</i>	competence protein CoiA-like family	3.9	Y	Y
<i>Locus VIII</i>					
kw2_1750	<i>comEC</i>	competence protein ComEC	3.1	Y	
kw2_1751	<i>comEA</i>	competence protein ComEA	3.1	Y	Y
<i>Locus IX</i>					

kw2_2122	<i>radA</i>	DNA repair protein RadA	11.9		Y
Locus X					
kw2_2177	<i>acdA</i>	zinc ABC transporter substrate-binding protein	2.0		
kw2_2178	<i>zitR</i>	zinc transport transcription regulator	1.8		
kw2_2179	<i>comGG</i>	competence protein ComGG	1.9	Y	
kw2_2180	<i>comGF</i>	competence protein ComGF	16.2		
kw2_2181	<i>comGE</i>	competence protein ComGE	6.9		
kw2_2182	<i>comGD</i>	competence protein ComGD	9.0	Y	
kw2_2183	<i>comGC</i>	competence protein ComGC	7.3	Y	
kw2_2184	<i>comGB</i>	competence protein ComGB	3.1	Y	
kw2_2185	<i>comGA</i>	competence protein ComGA	3.3	Y	Y
Locus XI					
kw2_2244	<i>recX</i>	recombination regulator RecX	1.9		Y
Competence associated genes					
kw2_1065	<i>tag</i>	DNA-3-methyladenine glycosylase Tag	2.3		
kw2_1849	<i>mecA</i>	MecA family protein	2.4		
kw2_0930		NlpC/P60 family protein	2.4		
kw2_0555		lytic murein transglycosylase	2.5		
kw2_1629		transglycosylase-like domain-containing protein	2.9		
Stress response and detoxication					
kw2_0055		OsmC family protein	2.3		
kw2_1396		glycine betaine/L-proline ABC transporter ATP-binding protein	2.4		
kw2_0182	<i>msrB</i>	methionine sulfoxide reductase B MsrB	2.5		
kw2_1287		ArsD family protein	2.7		
kw2_1288		ArsC family protein	2.6		
kw2_1322		toxic anion resistance protein	2.9		
kw2_1323		5-bromo-4-chloroindolyl phosphate hydrolysis protein	3.3		
kw2_1322		toxic anion resistance protein	2.9		
kw2_0306		AhpC/TSA family protein	2.4		
kw2_0340	<i>ahpC</i>	peroxiredoxin AhpC	2.6		
kw2_0388		thioredoxin family protein	2.3		
kw2_0410		superoxide dismutase	2.7		
Amino-acid metabolism					
kw2_2266	<i>glnA</i>	glutamine synthetase type I GlnA	2.3		
kw2_0565		methionine aminopeptidase	2.4		
Purine/pyrimidine biosynthesis and interconversion					
kw2_2006	<i>purA</i>	adenylosuccinate synthase PurA	3.1		

kw2_1464	<i>purE</i>	phosphoribosylaminoimidazole carboxylase catalytic subunit PurE	2.6
kw2_1476	<i>purM</i>	phosphoribosylformylglycinamide cyclo-ligase PurM	2.4
kw2_1091	<i>guaC</i>	guanosine monophosphate reductase GuaC	3.0

Transcriptional regulation

kw2_1424		transcriptional regulator MarR family	2.3
kw2_2267		transcriptional regulator MerR family	2.3
kw2_0564		transcriptional regulator TetR family	2.7
kw2_1642		two component system response regulator	2.6

Translation

kw2_1900	<i>infC</i>	translation initiation factor IF-3 InfC	2.3
kw2_0782	<i>rpsA</i>	ribosomal protein S1 RpsA	2.3
kw2_2212	<i>rpsB</i>	ribosomal protein S2 RpsB	2.4
kw2_1716	<i>rpsT</i>	ribosomal protein S20 RpsT	2.4
kw2_2056	<i>rplA</i>	ribosomal protein L1 RplA	2.3
kw2_2158	<i>rplB</i>	ribosomal protein L2 RplB	2.6
kw2_1225	<i>rplJ</i>	ribosomal protein L10 RplJ	2.5
kw2_2057	<i>rplK</i>	ribosomal protein L11 RplK	2.4
kw2_2335	<i>rplM</i>	ribosomal protein L13 RplM	2.3
kw2_2141	<i>rplO</i>	ribosomal protein L15 RplO	2.4
kw2_2159	<i>rplW</i>	ribosomal protein L23 RplW	2.4

Others

kw2_0426		sugar ABC transporter ATP-binding protein	2.3
kw2_0656		ABC transporter ATP-binding/permease protein	2.7
kw2_1072		cell surface protein	2.3
kw2_1381		basic membrane protein A	2.3
kw2_0534		enoyl-(acyl carrier protein) reductase	2.5
kw2_0861		FeS assembly SUF system protein	2.3
kw2_1842	<i>sufB</i>	FeS assembly protein SufB	2.3
kw2_0839	<i>mraW</i>	S-adenosyl-methyltransferase MraW	2.6
kw2_1319	<i>pyk</i>	pyruvate kinase Pyk	2.5
kw2_0318		ybaK/ebcC protein	2.5
kw2_0472		DegV family protein	2.5
kw2_1145		aldo/keto reductase family protein	2.3
kw2_1228		HTH/peptidase S24 family domain- containing protein	2.3
kw2_1984		nitroreductase family protein	2.8
kw2_2173		FMN-binding protein	2.3

kw2_0880	glyoxalase family protein	2.5
kw2_1964	PspC domain-containing protein	3.9
kw2_1965	PspC domain-containing protein	4.6
Hypothetical proteins		
kw2_0067	hypothetical protein	3.2
kw2_0305	hypothetical protein	3.0
kw2_0315	hypothetical protein	3.1
kw2_0615	hypothetical protein	2.6
kw2_0862	hypothetical protein	3.2
kw2_0934	hypothetical protein	2.3
kw2_0971	hypothetical protein	2.4
kw2_1019	hypothetical protein	2.4
kw2_1025	hypothetical protein	2.9
kw2_1713	hypothetical protein	2.4
kw2_1913	hypothetical protein	2.3
kw2_1966	hypothetical protein	5.1
kw2_2070	hypothetical protein	2.3

^a Genes were considered as up-regulated when the calculated fold change (FC) was ≥ 2.0 with DESeq2, edgeR and DEXUS packages. The FC value is the mean of the FC obtained with the 3 packages. When the up-regulated genes were preceded by a predicted ComX-binding motif (Com-box), adjacent downstream genes with a calculated FC ≥ 1.5 with the 3 packages were also retained.

^b Orthologous late *com* genes of *S. pneumoniae* shown as essential for DNA transformation (1); Y, presence.

^c Predicted Com-box recognized by ComX (Table S3); Y, presence.

TABLE S3. Predicted Com-box in intergenic regions of the genome of KW2.

Locus tag	Gene	Description	Up-regulated ^a	P value ^b	start ^c	Predicted Com-box ^d
kw2_0357	<i>recA</i>	recombination protein A RecA	Y	1.3E-07	-78	GTGACAAAGTTCATAAAAAAACGTATA
kw2_0391	<i>ssbB</i>	single-strand binding protein SsbB	Y	1.1E-11	-21	GTGACAAATGCTCAAACTTTACGTATA
kw2_0833		hemolysin TlyA family protein		4.9E-08	-16	GATACATTTTTTAGAAAAATACGTATA
kw2_1003	<i>radC</i>	DNA repair protein RadC	Y	1.2E-11	-7	GTGACAAACGAATGTGTTTTTCGTATA
kw2_1034	<i>comFA</i>	competence protein ComFA	Y	1.6E-11	-7	GTTACATTTGAACAGCATTTCCGTATA
kw2_1192	<i>dprA</i>	DNA protecting protein DprA	Y	2.8E-10	-61	GTTACAATTTAAATATTTTATCGTATA
kw2_1577		cell surface protein	Y	5.7E-08	-19	GTTACAAAATCTCCTTTTTCGTATA
kw2_1705	<i>coiA</i>	competence protein CoiA-like family	Y	6.8E-10	-7	GTGACAAAATTATATTTTTCGTATA
kw2_1751	<i>comEA</i>	competence protein ComEA	Y	1.1E-08	-7	GAACTTTTCAAGAGTTTTCGTATA
kw2_1863	<i>recQ</i>	ATP-dependent DNA helicase RecQ		3.6E-06	-11	GTGATTTTTCCCTCTCTTTACGTATA
kw2_2080	<i>comC</i>	type IV prepilin peptidase		6.0E-09	-7	GTTACATTCAAAAICCATTTTCGTATA
kw2_2122	<i>radA</i>	DNA repair protein RadA	Y	1.3E-11	-16	GTTACATCTGAATAAATTTTACGTATA
kw2_2185	<i>comGA</i>	competence protein ComGA	Y	4.6E-09	-7	GTTACATTTACTAACAAATTCGTATA
kw2_2244	<i>recX</i>	recombination regulator RecX	Y	1.1E-08	-10	GTGACAAATTCCTAAGATTTTCGTATA
kw2_2251		6-phosphogluconate dehydrogenase-like protein		1.1E-06	-43	TTGACAAATGAAAAAAATTCCTATA

^a Up-regulated genes (Y) according to Table S2.

^b P value < 1.0 E-06.

^c Position of the start of the predicted Com-box relative to the first nucleotide of the start codon of the downstream gene.

^d Predicted Com-boxes identified by MEME and MAST (<http://meme-suite.org/>) using as input the Com-boxes upstream of *comEA*, *comFA*, *comGA*, *dprA*, *coiA*, *ssbB*, *radA*, *radC*, and *recA*.

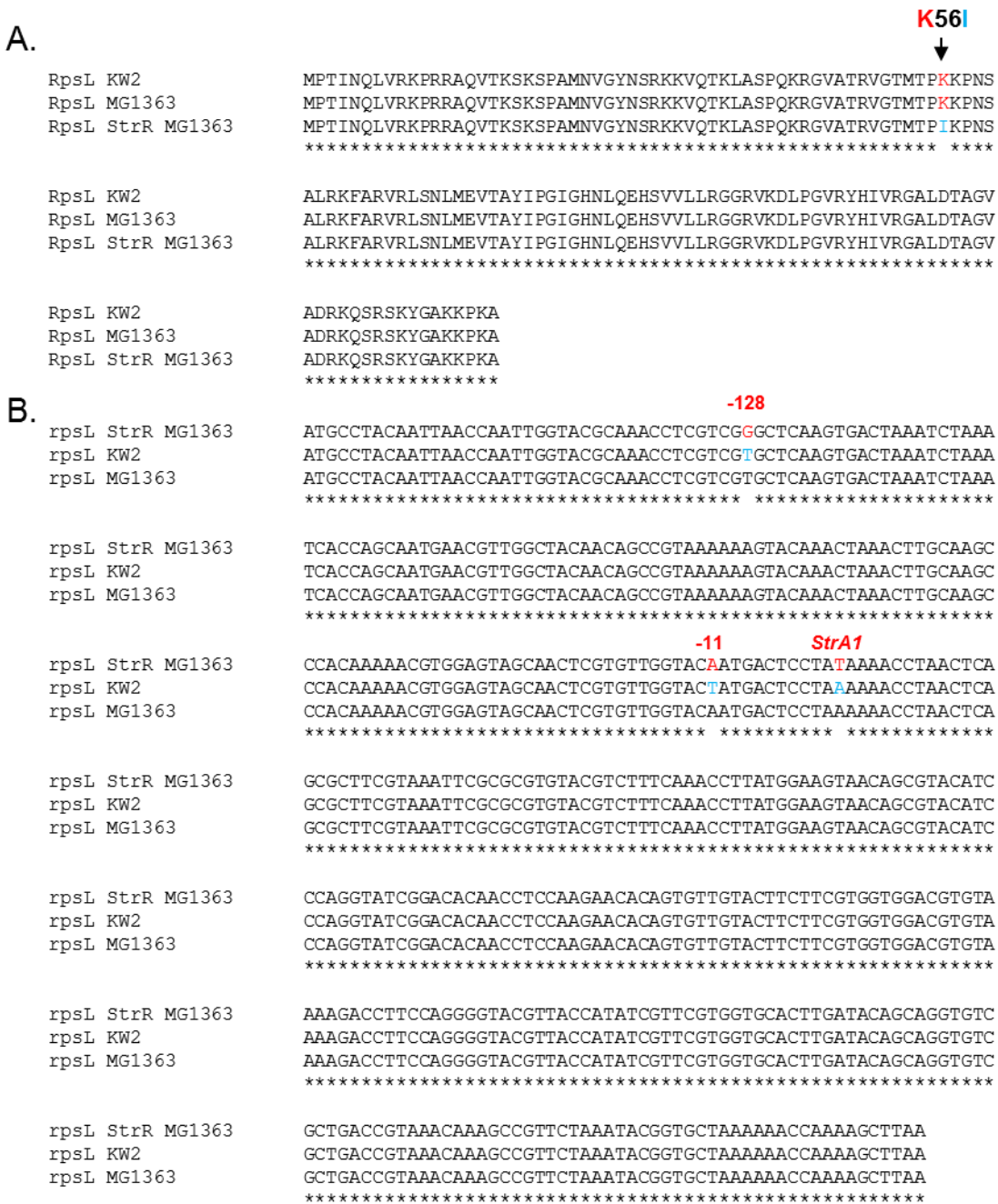


FIG. S1. Protein and nucleotide mutations of a selected spontaneous *Str^r* mutant of MG1363. A. Alignment of ribosomal proteins S12 (RpsL) from KW2, MG1363 and the selected spontaneous *Str^r* mutant of MG1363. The K56I mutation (*strA1* allele) is highlighted. B. Nucleotide alignment of *rpsL* ORFs from KW2, MG1363 and the selected spontaneous *Str^r* mutant of MG1363. The three mutations (*strA1*, -11 and -128) in the nucleotide sequence of the mutated *rpsL* variant of MG1363 compared to *rpsL* of KW2 are highlighted.

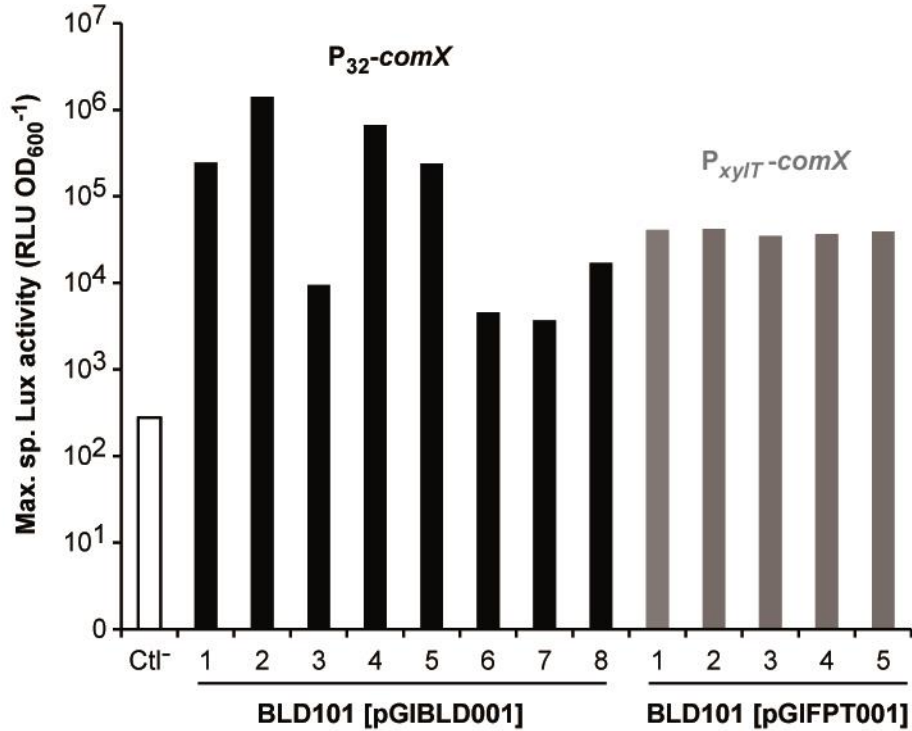


FIG. S2. Stabilization of the activation of the late promoter P_{comGA} by inducible *comX* expression. Maximum specific luciferase (Lux) activity (RLU OD₆₀₀⁻¹) emitted by 8 independent clones (clone 1 to 8) of BLD101 [pGIBLD001] (P_{32-comX}) compared to 5 independent clones (clone 1 to 5) of BLD101 [pGIFPT001] (P_{xyIT-comX}). Ctl⁻, control strain carrying the empty vector (BLD101 [pG⁺host9]). Cells of BLD101 [pGIBLD001] and BLD101 [pGIFPT001] were respectively grown in CDMG and M17X. Data were collected from Fig. 2A and 5B.

REFERENCES

1. **Peterson, S. N., C. K. Sung, R. Cline, B. V. Desai, E. C. Snesrud, P. Luo, J. Walling, H. Li, M. Mintz, G. Tsegaye, P. C. Burr, Y. Do, S. Ahn, J. Gilbert, R. D. Fleischmann, and D. A. Morrison.** 2004. Identification of competence pheromone responsive genes in *Streptococcus pneumoniae* by use of DNA microarrays. *Mol. Microbiol.* 51:1051-1070.