

***Sinorhizobium meliloti* GR4 produces chromosomal- and pSymA-encoded Type IVc Pili that influence the interaction with alfalfa plants**

Cristina Carvia-Hermoso¹, Virginia Cuéllar¹, Lydia M. Bernabéu-Roda¹, Pieter van Dillewijn^{1,*}, María José Soto^{1,*}

List of Supplementary Material

Figure S1. Predictive protein composition and structure of the secretion systems encoded by *S. meliloti* *tad/cpa* genes of the *flp-1* cluster, *flp-2* cluster and Locus 3.

Figure S2. Multiple sequence alignment by MUSCLE of Flp prepilins.

Figure S3. Multiple sequence alignment by MUSCLE of PilQ2-like gene products.

Figure S4. Multiple sequence alignment by MUSCLE of SMa0164 with homologous proteins of *S. fredii* NGR234 (Sf) and *B. diazoefficiens* (Bd).

Figure S5. Multiple sequence alignment by MUSCLE of SMa0166 with TadD-like gene products.

Figure S6. Transmission electron microscope (TEM) images of flagella-plus single and double *flp* deletion mutants.

Figure S7. Twitching motility assay of duplicate stab inoculations of the different strains on different media with different agar concentrations and types.

Figure S8. Appearance of alfalfa plants inoculated with *S. meliloti flp* mutant strains at the end of the nodulation kinetics experiment.

Table S1. Percent sequence identities (in parenthesis) among *S. meliloti* T4cP paralogues.

Table S2. Strains and plasmids used in this study.

Table S3. List of primers used in this study.

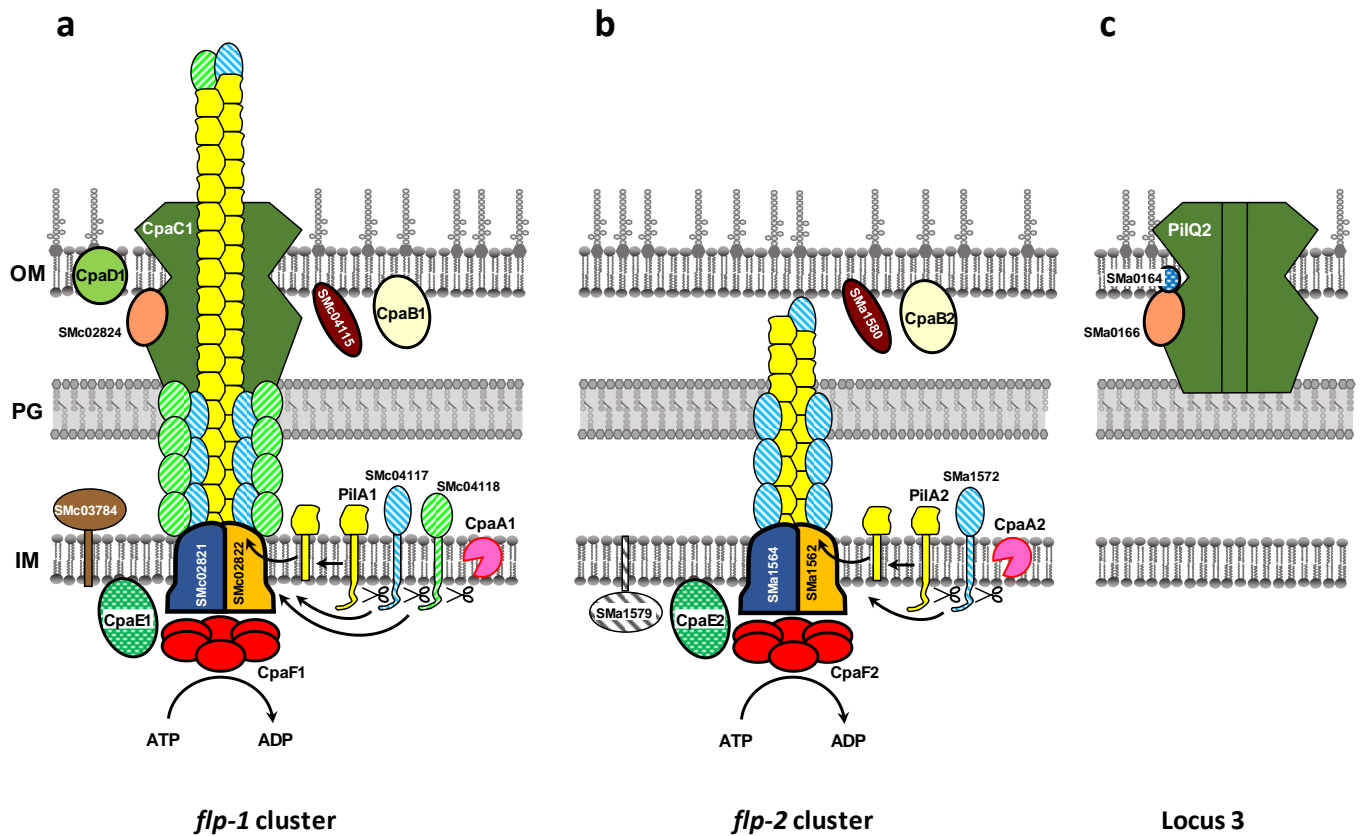
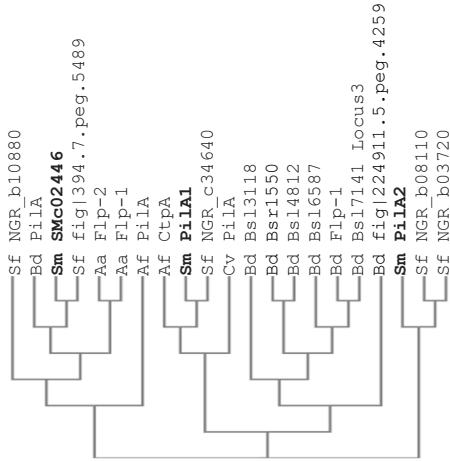


Figure S1. Predictive protein composition and structure of the secretion systems encoded by *S. meliloti* *tad/cpa* genes of the *flp-1* cluster (a), *flp-2* cluster (b) and Locus 3 (c). See text in the main manuscript for details.



Sm	SmC02446	-METTRRLRDHGATAVEGLLAAALISVGLLTGLQNFSCALLGMLFVNTVITAAWIT--
Sm	PilA1	MKTIIFARLKDSEGATAIDYGTAAALISVALIGQAQTLGGASQFTNIGLYLNVENPAP
Sm	PilA2	MKNLILFARNESGATAIDYGLTAGLISVLLITVMGTITGLTTRFTATGTALTGG----

---MKAFTVKVSFAREEDGVALSEYLVJLGLL---VGVIAAV---TTVSGN-LAAATWSGUTFTSHLSCDPAALPL
 MURITPLWSTRNPPSEHYKRIQJQAQGFANEHSEVCHLQGRDPPMLTVYIETREALKRLRTDQGVVSFEYIIIVAVCIYAGAVGGGAGGTGAA-LTGITAIITAFATAFATIG---
 ---METLRLRLRDHGATAVEYGLIAALI---SVGLIATGL---QNFSGA-LTGITAIITFVNTTEAAWI---
 ---MDKLRSIVRNRDGAATVEYGLIAALI---SVALIVGL---QGFSE-LNVNLIIADAEDSWT---
 ---MMDLIDYFYRQVFFSCNFYRNRQGTISVEYGLIAAV---AVFVAVLVDGNSFIKA-LSGKFDLTTTSGVAILVSKSS---
 ---MLNTLITTKAYIKASEAIRSFRENOAGVTAIEYGLIAFV---AVLIVAVFYSNNGFIAN-LOGKFNSLASTVASANVTK---
 ---NPVVBETHRMVDEYFVPLHFFINFCNKNENGATAIEYGLIAGII---SAAIIAGL---GNISG-INAVQFOIVDAFPKG---
 ---MTKIFARFLKDESGATAIEYGLIAALI---SVAIIGGA---STLGGK-LKDFTFIKGSFTDKRASTGGV---
 ---MKTIFARLMKDESGATAIEYGLIAALI---SVALIGGA---QTLGGA-LSTQFTNLGGLYNVEPNAP---
 ---MKTIFARLMKDESGATAIEYGLIAALI---SVALITGA---TALGDS-LOSMFNALSQOMTAEITNM---
 ---MTKFVTRFLKDESGATAIEYGLIIVALI---AVVIVTAV---TTLGTN-LRTAFTKAGAAVSTAAGT---
 ---MALLKSFLEDSATAIEYGLIAGI---ALVIVTVW---NNTGSALLNNKFNSIDAATK---
 ---MKTIVHFLRDESGATAIEYGLIAGI---SLAIIAVW---NLGTR-LNTKFSISSLK---
 ---MICEHACQRYRGATMKNTLKNFLADERGATAIEYGLIAGI---ALAVITVW---NGWGSK-LNTKFGSISTSLK---
 ---MTRMILKFSWNSGATAIEYGLIAGI---ALAIITVW---NLGTT-MNEKFTSISTSLK---
 ---MKNLILARFAKDESGATAIEYGLIAGI---ALAIITVW---NNLGST-LNTKFTSISTSLK---
 ---MVGRRWEAEETRPVVSCTFWEHSMKNILVRFVKDESGATAIEYGLIAGI---ALAIITVW---NNLGST-LNTKFTSISTSLK---
 ---MKLICLLADFADESGATAIEYGLIAGI---ALAIIEII---YALGTN-LVAKLOSALTALK---
 ---MKNLILARFARNESGATAIEYGLIAGI---SVVLITVW---GTIGTR-LITRTAIGTATFGG---
 ---MKNLILLRFRVHESGATAIEYGLITGLI---AV--ITAV---QTVGTD-IDAAPDAIAAL---
 ---MKNLILRFRVNESGATAIEYGLIAGI---AVVIISAV---QIVGTD-I GAKFTAI STAL---

Figure S2. Multiple sequence alignment by MUSCLE of Flp prepilins. **(a)** Alignment of the three putative Flp prepilins of *Sinorhizobium meliloti* (Sm). **(b)** Alignment of the Flp prepilins of *S. meliloti* with known Flp/PilA/CtpA prepilins of *Aggregatibacter actinomycetemcomitans* (Aa), *Caulobacter vibrioides* (Cv), *Agrobacterium fabrum* (Af), putative prepilins of *Sinorhizobium fredii* NGR234 (Sf), and *Bradyrhizobium diazoefficiens* (Bd). Identical residues are indicated with an asterisk. Colons and full stops indicate partial matches. The Flp motif is highlighted in grey. In **(b)** the cladogram obtained in MUSCLE is also shown.

Aa	RcpA	-----MNNWHITFGKQLICCAVLGAVFSL-----	-----NAYAQ Q MSLDKGAQLVQVTKKIKIDITFVSSPNIAIDVILIDNITFIYAKBEGRETEVTATGADGRPTSDTVVND
Cv	CpAc	-----MSREVNLSVKASMSAGSRLLITASVALLAITSTAPVFVDDGIIGSGSHYRPPARVYTPAPAAPVPMVPAEDQVARTVMVTAQ Q AAITLLELAKKSAALVELPEVVDLIVTPQOIAADVATDPRKRIYIVGLVAGQNTDAAFFTAGRRLLSLSLIRV	-----GNGIRKHLFGLKALVLELPEDAADILVADPILADATFTRSRITFIPGRVWGQNTFIIFGGREELVSLDLE
Af	CtpD	-----MTGLGLMTKRFKHLHRSMSAGGLAFLCTFS-----	-----GEG Q AGTINLGIKAKAVVLELPEDAAILVADPILADATFTRSRITFIPKSVGQNTFIIFVGGEEVSLSLVANE
Sm	CpAc1	-----MKPFAKKFRASILAAGLSFCLTFS-----	-----GEG Q AGTINLGIKAKAVVLELPEDAAILVADPILADATFTRSRITFIPKSVGQNTFIIFVGGEEVSLSLVANE
Sf	NGR_c34670	-----MKRSQNTFRCLLAAGLSFSLTFS-----	-----GGCVKRWLQDQKNAVVVLELPTDAADILVADPILADATFTRSRITFIPKSVGQNTFIIFVGGEEVSLSLVANE
Bd	RcpA	-----MSYDMKKRADLMTWRTSMVRLSFAALA-----	-----GQMARELSLGVKSLVITLPRDIDQVIVVADPKIANVATERSQVATIGAAVQNTNIVVFTISAPQOAAAYDIIVK
Bd	B110666	-----MNYGDDRAGLRIRGNRARTFWAGAMIT-----	-----ENNDYQRVAVNKKSRITFKYDITAFATIIAGSEFIDVMSKLSLDRHLYVQKQGTGTTNVLIFESAMQIGVLDIVK
Bd	CpAc	-----MGGGRVSGKTVARLLAFGVMGLGAVALIC-----	-----LJGSLVQVTLAINDTITVPRGKPDQVITGSAEMINIVPELSDOSLFTBKQQTGAIMI SVYDMKRALNGVLDIVRA
Sf	NGR_b10860	-----MSKGLIKCAARACTLAALAAVLVV-----	-----SSN S SESLVAKGKVITFVAGV Q FEVILVIGIEVIAVTEFILTUSEVYILNVLISLSESVNITFPAEKRGVIGLIDIEV
Sm	PI102	-----MIRSNGSLIFAFCEVVALIC-----	-----
Aa	RcpA	-----SVWTSIADINKQLKSRPENTNLSVKVKGAVIEGKARQOESDEVRIIVGEALGSRKVTETKLG-----	-----EDSLPELDKVHYDGVWDNANIDTITQINVKLSWEVNNKLSAEMGNINSHWAGS--G-----
Cv	CpAc	-----OPVDOLAAM-----	-----KV LA MI SVA KGQDWMLQVRLVEVORVITKQLGVDAINAVTIGOLGETOYTFQAPGYGNSGLLGVTGGVK
Af	CtpD	-----RD EGLEAN -----	-----TTCGNNMGDAITFAERQVSOIVNMILTTEGEDQVILKTVAEVSRQVTKOLGENGS--SSSNG--
Sm	CpAc1	-----RDA AGLEAN -----	-----GDADIFAEIQNSQYNLITTEGGDQVILKTVAEVSRQVTKOLGEN--GRISD-G-----
Sf	NGR_c34670	-----RD AGLEAN -----	-----GDAADIFAEETQSNVLLITVEGEDQVILKTVAEVSRQVTKOLGEN--GRISD-G-----
Bd	RcpA	-----RDIN GLEAN -----	-----KVANSYVGRQDWMLKTVAEVARSIVKQLGIDITANLNY--G-----
Bd	B110666	-----RDINGMRAA-----	-----KVANSYVGRQDWMLKTVAEVARSIVKQLGVLSASINA--G-----
Bd	CpAc	-----IDCN LOQ -----	-----KVANSYVGRQDWMLKTVAEVARSIVKQLGVLSAANAN--GTNIGNSGLGRVTAAGRTPIGGINT--
Sf	NGR_b10860	-----ADPSEVASA-----	-----V AN LS KTU SSQVQVLNVRVIRNDRVGRKLQGISAAW-S-----
Sm	PI102	-----LDT ILSS -----	-----
Aa	RcpA	-----MDTTKOPVMTPCTGCGMTAGSLCPVARGGNSDT-----	-----SISAFINALDQNSNGKVLAEPNISMISGETADILVGGIEIPFAQRDR-----
Cv	CpAc	-----ANPS-----	-----SAKGMQAFERGLVRLTAEPNIAAYSGESGKFLVGGEPVETGSDT-----
Sm	CpAc1	-----RNPA-----	-----RFSYINAMEQGVNRLTAEPNIAAISGDAQKYVGGDFRILAAQOEVIDKE-----
Sf	NGR_c34670	-----RNPA-----	-----TISSYNMEQGVNRLTAEPNIAAISGDAQKYVGGDFRILAAQOEVIDKE-----
Bd	RcpA	-----Bd	-----SVQATIRAMEVAGVIRLTAEPNIAISGESATFIAGGEPFVAGYACDPTTH--VCTT-----
Bd	B110666	-----Bd	-----TWNATIRAMEVAGVIRLTAEPNIAISGESATFIAGGEPFVAGYACDPTTH--VCTT-----
Bd	CpAc	-----Bd	-----NGPP-----
Sf	NGR_b10860	-----Bd	-----NGAT-----
Sm	PI102	-----Bd	-----DGTI-----
Aa	RcpA	-----KLOKNRRLALDQVSTIAGNVNFDGIGN-----	-----IPFENTRKAISLFEVANGESFIIIGGLFSSNDLEGINVKVPIIGDIPLIGSFFRSATTERKKELVIAVNVVKPVNE--KDVIVPD--TEKTGTMERFFHTPLKNVY----YKT
Cv	CpAc	-----VVISGGRTSLKISTVSELSLIGAFS--MSCTSSGSTN-----	-----ITVPGLSVRVSTIVSELPSGSILMAGLLOQTKETLDSLPGMSNPLGLSPRSDFLANQDELWILITPVILDPTPE--QNLCQPADGLQFASDQMTVILGR--LNKVKYKAPAGAN
Af	CtpD	-----VVISPGRTSLKIETWSEFTYEGNVV--TNGGGRTIPGS-----	-----TMSVRRARETSTIVSELPSGSIVTLAGLVQDNVQMSGLPGMSKIPITGTLIFRSKEFLNETELWILITPVILVRPVAR--NQIARPDNDNFENDAAFMNR--LNKVKYKQDQOQ
Sm	CpAc1	-----VVISPGRTSLKIETWSEFTYEGSVV--TGNMSTAIPGN-----	-----TELGIRREASTSVSELPSGSIVTLAGLVQDNVQMSGLPGMSKIPITGTLIFRSKDFQSESELWILITPVILVRPVAR--SALSRRPDNDNFENDLESFFLGR--VNRILYGRPEAP
Sf	NGR_c34670	-----VVISAGRTSLAIETVSEPTYEGSVV--TGNMSTFDPGN-----	-----VVISAGRTSLAIETVSEPTYEGSVV--TGNMSTFDPGN--PILGTLIFRSDFORNESESELWILITPVILVRPVAR--SALSRRPDNDNFENDLESFFLGR--VNRILYGRPEAP
Bd	RcpA	-----VVISAGRTSLAIETVSEPTYEGSVV--TGNMSTFDPGN-----	-----VVISAGRTSLAIETVSEPTYEGSVV--TGNMSTFDPGN--PILGTLIFRSDFORNESESELWILITPVILVRPVAR--SALSRRPDNDNFENDLESFFLGR--VNRILYGRPEAP
Bd	B110666	-----VVISAGRTSLAIETVSEPTYEGSVV--TGNMSTFDPGN-----	-----VVISAGRTSLAIETVSEPTYEGSVV--TGNMSTFDPGN--PILGTLIFRSDFORNESESELWILITPVILVRPVAR--SALSRRPDNDNFENDLESFFLGR--VNRILYGRPEAP
Bd	CpAc	-----VVISAGRTSLAIETVSEPTYEGSVV--TGNMSTFDPGN-----	-----VVISAGRTSLAIETVSEPTYEGSVV--TGNMSTFDPGN--PILGTLIFRSDFORNESESELWILITPVILVRPVAR--SALSRRPDNDNFENDLESFFLGR--VNRILYGRPEAP
Sf	NGR_b10860	-----VVISAGRTSLAIETVSEPTYEGSVV--TGNMSTFDPGN-----	-----VVISAGRTSLAIETVSEPTYEGSVV--TGNMSTFDPGN--PILGTLIFRSDFORNESESELWILITPVILVRPVAR--SALSRRPDNDNFENDLESFFLGR--VNRILYGRPEAP
Sm	PI102	-----VVISAGRTSLAIETVSEPTYEGSVV--TGNMSTFDPGN-----	-----VVISAGRTSLAIETVSEPTYEGSVV--TGNMSTFDPGN--PILGTLIFRSDFORNESESELWILITPVILVRPVAR--SALSRRPDNDNFENDLESFFLGR--VNRILYGRPEAP
Aa	RcpA	-----LTSNFI-----	-----KNSGFIQ-----
Cv	CpAc	-----AGRAYQ-----	-----GPVGVIIE-----
Af	CtpD	-----AA-PYQ-----	-----GSVGTIYK-----
Sm	CpAc1	-----PAGRYH-----	-----GNVGFTYK-----
Sf	NGR_c34670	-----PVRGYH-----	-----GNVGFTYK-----
Bd	RcpA	-----PGNRYR-----	-----GTYGFITD-----
Bd	B110666	-----PIDGAR-----	-----GDFGITD-----
Bd	CpAc	-----SGEDVK-----	-----GPYGHILIAETVETVFPFAANQFVVKTN
Sf	NGR_b10860	-----NGEDVK-----	-----GPYGHILIAETVETVFPFAANQFVVKTK
Sm	PI102	-----ASRQARVVRAPSSGHFLEIQ-----	-----

Figure S3. Multiple sequence alignment by MUSCLE of PilQ2-like gene products. Alignment of the amino acid sequences of PilQ2 (SMa0163 Locus 3) of *S. meliloti* with PilQ2-like gene products of *A. actinomycetemcomitans* (Aa), *C. vibrioides* (Cv), *A. fabrum* (Af), *S. fredii* NGR234 (Sf) and *B. diazoefficiens* (Bd). Identical residues are indicated with an asterisk. Colons and full stops indicate partial matches. The Signal peptide motif (Phobius) is highlighted in yellow, the T2SS-T3SS pil N: Pilus formation protein, N-terminal domain (InterPro IPR032789) in green, the BON domain (Pfam PF04975) in red and the T2SS/T3SS Dom: Type II/III secretion system, secretin-like domain (IPR004846 InterPro) in grey.

```

Sf NGR b10850      MSF PRSKRAI---LVLTAAGILSGC-----ADYLNHRDSITFGLGNAVEANKGIH
Sm Sma0164         -MIRRS SRPRGF PGLLFASSIILSGC-----QNHVLRSETIALSAGDAIAANSVMQ
Bd blr3501         -MTKRY-----LVLFAPLILGGC YGLAGNDEM DRYFQRSNTITMSAGDAKEVNAVTH
                   *      *::: . . **.*
Sf NGR b10850      IQDPFPRAAYNTQIASDGKVIHRAIRTYQGSQAQTSAP-----
Sm Sma0164         MVD PWP PRVKQTSLATPAD----LEQYKPQQPNAEQ-----
Bd blr3501         TIHPWPRNVGDRRIAYDARRVGAAVTRYGKTQQPVDQLPDISDPTRQMGVRTTPTQNVNI
                   *: * . : : * . : *
Sf NGR b10850      -----SPSAVILPVATP PNGAAR
Sm Sma0164         -----NGGNGETYPNDTTTQ----
Bd blr3501         EGLGAGSSTGVSVPVGGTGGAPNR
                   . . . * .

```

Figure S4. Multiple sequence alignment by MUSCLE of Sma0164 with homologous proteins of *S. fredii* NGR234 (Sf) and *B. diazoefficiens* (Bd). Identical residues are indicated with an asterisk. Colons and full stops indicate partial matches. The prokaryotic membrane lipoprotein lipid attachment site profile is highlighted in yellow. The lipobox motif (LXXC) is indicated with a box.

Aa	TadD	----- NY SKSLKAVLLGGHLSVTA SA VINKKFLSAD-----KLAAREALYOSTNNIDALITHYDGLKWKEDPLTR-----YKLSIIYKKGDSNSSLIIYLOPLL-----TNS
Bd	TadD Locus2	KSPMSRAGT TVRRITIIIMLYCCIGTSLAG CD YITLRAAYAP-----VDPEGG-----DPVOEFTDVKY-----YESDE
Sf	NGR_b10740	-----MIDQLOSRGLSPILLAAVLVIGGOSAAFDVAFG-----DSAKTLEDOSAVAFY-----DPTSALASSGDSV-----KN
Sm	SMa0166	-----HEMTGASGRSV-----LGRSGGRAPTFLA-----MAIVATV LS CCQTSEVLSGAEF-----DPTSALASSGDSV-----
Cv	CpaO	-----MCRGALIAVLAPTAIVLTAIVPFAAPFFPAADAGTKAKAPAFVKAASQORTFAARRDPLAQAFWGAEFVDDADAFAGAGLAKAFALGRADFAAIAATKGLIAHPDOT
Af	Atu0215	-----MTAFSLGETRTGILRGACLAALVAVSGCASTNEDPMSTIG-----SIPAAITRSTYHLNMQLQAEETAGKAVRNPKDSVQNTASVLMTKNTQALAVMQQIAIANPADR
Sm	SMc02824	-----MTMQINFESLRRLIATGVAVAMVALISISAGAGCCGRKELITIG-----SIEPLTRFWQSMNATEIAAAAEETIGQAVRNPKDREAGLNYANLLRMTGRNEQALAVMQQVAIVHPADR
Sf	NGR_c34740	-----MRLAVGAAVALMALALGCGGAGCCGRKELITIG-----SIEPLSKFVQSMNATEIAAAAEENIGQAVRDPKREAGLNYANLLRMTGRNEQALAVMQQVAIVHPADR
Bd	TadD Locus1	----- MQRFPSPARILASLSLVAVVAVMSLGG TAMSKLSDV-TG-----SIGSRA-----EASPPSDPAGAVFYGYERYRANPKDAEALGYQALRANGQARAARVLEQATIANFENK
Bd	Blr0663	----- MSKRILSLASRPA RVL SA LVLLALGGCQT-TGIEDI-TG-----ALGSKS-----EAAASKADGKPTMDALRRERYRAKPSDENVALDYGKALRESQSAQAVAVLEQAVLGHFENK
Aa	TadD	-----GQLSEKAVLQARNINOLKRYOEALFVENSLLITSPKNEVYNLRGVTVLMMGNENKANEINRAREYFLNDVAVNNILAMLSIVNGDYRNNAVSVLLLPQYLNQVREQRLVHNLIIFALIKN
Bd	TadD Locus2	-----FV--RLGLEHFNRG-----SYGLSQRYFKDAVEKSPKDLTAMIGLAASYDRURRFDLAQAAQAQIRLGGETVOILNDQGYSTMLRGNTLSAARRKFEKAYSIDPQNFVIANNLEILNGS--
Sf	NGR_b10740	-----DELITTTGLOFKEK-----NYGKSYALYKRAVANVFQDPSAWLGFASADMIGRFEDTSDRAVQOLSKMTIGNSPVYNNNGYSRLRGDLPRAROYFLKAYELDPANETTARNLEIM----
Sm	SMa0166	-----KSDLDQGGKQFMNG-----NYGLAEGRFRKAVELRQDNAAALMGLAACYDRLGRFDLADRAINQLLVAGROPRIVNNMGYSQVLRGEEKAKARKLILLEAPASFGNETIEINAIL----
Cv	CpaO	-----GLLIELARAHIAHG-----QGYTAIEEPARKVAALSPKDWRLPLITLLGVAYEQAEEDFEAAHRQALIAAPNEATPMANHAMHLAARGDLAGAEELRRATVLPASGTQVRQNIALAVVGLQ
Af	Atu0215	-----EVLAAYGKSOQAAG-----QLEQALSTIQRAQTPDRPDWRLYSAE GAVLDQGRSNEARSRVROALDKPNDPSVLNIGMSYLLSDSDPTAETVLOSALISQFGADSRVRQNIALAVVGLQ
Sm	SMc02824	-----EVLGAYGKAOQAAG-----QLEQALATISRAQTPDRPDWRLKSAEGAILDQGRSARLRVREADLKPENEPSVLNIGMSYLLTKDLRTAETVYLSAASQFDAGSVRQNIALAVVGLQ
Sf	NGR_c34740	-----EVLGAYGKAOQAAG-----QLEQALQTIIGRAQTPDRPDWRLKSAEGAVLDQGRSSDARLRVREADLKQNEPTLLNIGMSYLLTKDLRTAETVYLSAASQFDAGSVRQNIALAVVGLQ
Bd	TadD Locus1	-----ALLALYGRALADNG-----NFQAFDVLSKAHSEPDNPDWRLISVQGTALDQGRHDEARSYYASALKIAPGDGPVLSNIGLSYLLSRDLPKAEELRQAYASPRASRVQRNLGIIVVGLQ
Bd	Blr0663	-----ALLAGYGRALADNG-----NFQAFDVLSRAHTEPDPDWRLISQAALDQGRNEEAQQYVATALKIVYPEPOVLNIGLSYLLQNLLPKAEQVLGRAHQNRQNIDARIRANIALVGLQ
Aa	TadD	-----NDIDYAKDIIYKENINTSPDDLINALKKTERTVSNNIAR-----
Bd	TadD Locus2	-----RRFIERPENNOF-----
Sf	NGR_b10740	-----KNSVYKVAQR-----
Sm	SMa0166	-----DRS-----
Cv	CpaO	-----GRLFEAEKLVRADLPPEQVANNLAYIRAAVGGQSRWDAMRAGGGR-----
Af	Atu0215	-----GRFAEAEIRAVOELSAQQAQANLTYLRAMLSQNSMQOLAKMDKRPAG-----
Sm	SMc02824	-----GRFOEATIIAQOELTSEQEANVAYLRVLSQCGANWQOLAKAD-----
Sf	NGR_c34740	-----GRFOEATIIAQOELTPEQEAANVAYLRKMLSGKGMATELAKADDGNTN-----
Bd	TadD Locus1	-----GRFAEATIVKADLPPEQAAANVAYLKEMLSRSDAPGAPKRTFVAALSRPD-----
Bd	Blr0663	-----GREAEAEILIVKADLPPEQAAAAKITAIRLOLLAKKQOQRAEK-----

Figure S5. Multiple sequence alignment by MUSCLE of Sma0166 with TadD-like gene products. Alignment of the amino acid sequences of Sma0166 (Locus 3) of *S. meliloti* with TadD-like gene products of *A. actinomycetemcomitans* (Aa), *C. vibrioides* (Cv), *A. fabrum* (Af), *S. fredii* NGR234 (Sf) and *B. diazoefficiens* (Bd). Identical residues are indicated with an asterisk. Colons and full stops indicate partial matches. The Signal peptide motif (Phobius) is highlighted in yellow and the Tetrapeptide-like helical domain superfamily (TPR-motif) (InterPro IPR011990) is highlighted in grey. The prokaryotic membrane lipoprotein lipid attachment site profile (PROSITE PS51257) and the lipobox motif in Sma0166 are indicated in red letters.

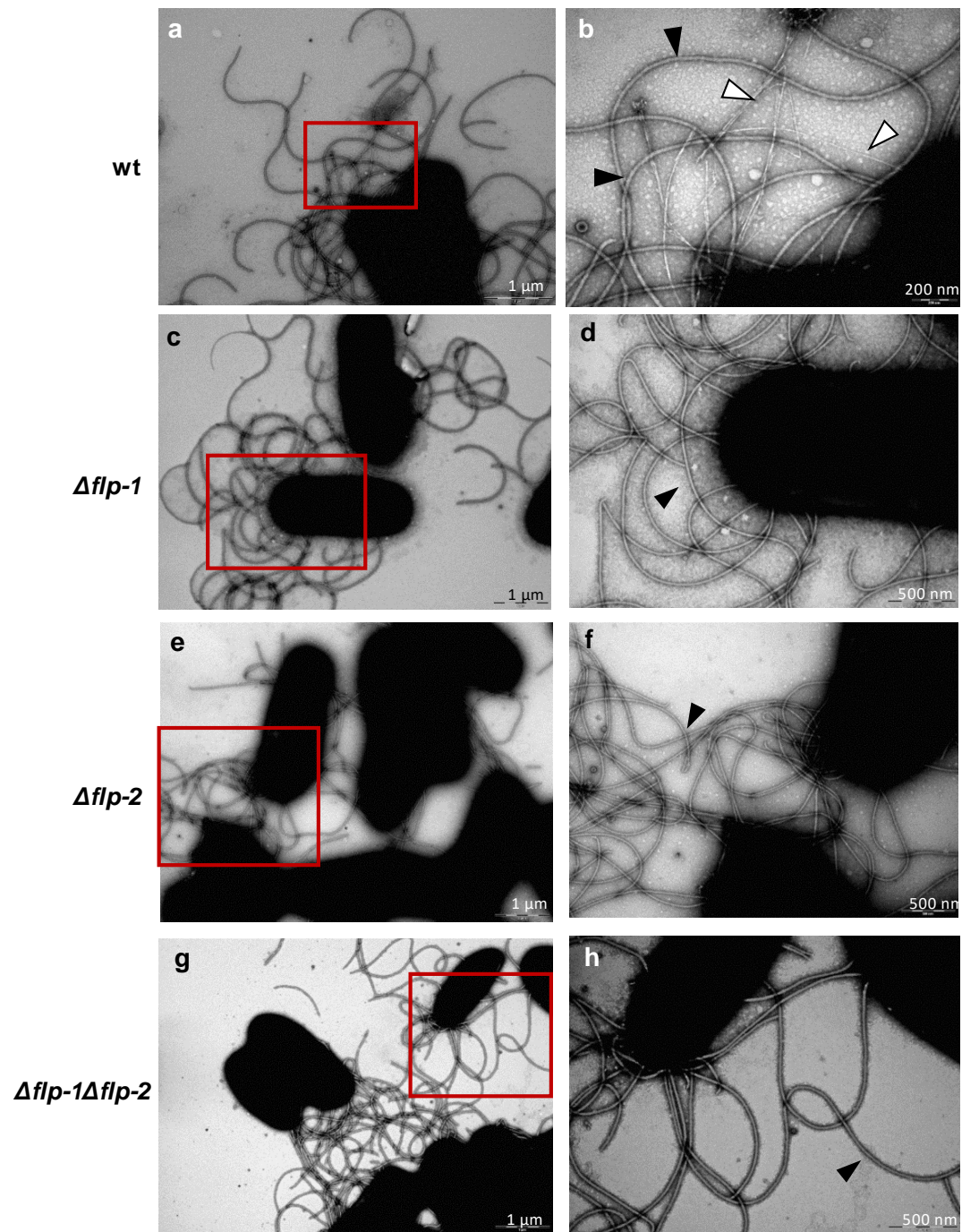


Figure S6. Transmission electron microscope (TEM) images of flagella-plus single and double *flp* deletion mutants. Cells were grown on the surface of Minimal Medium (1%) and negatively stained with uranyl acetate (2%). Red squares indicate the zone shown at higher magnification on the right. Solid and open arrowheads indicate flagella and pili-like structures, respectively.

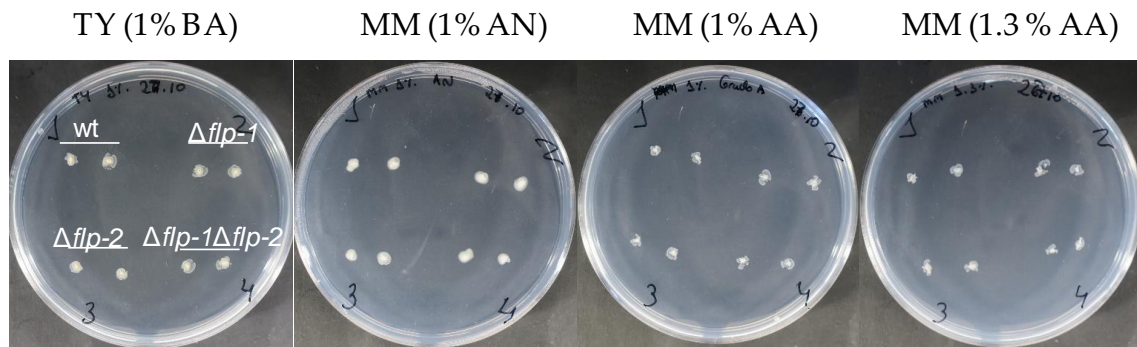


Figure S7. Twitching motility assay of duplicate stab inoculations of the different strains on different media with different agar concentrations and types. BA: Bacteriologic Agar; AN: Agar, Noble; AA: Agar, Grade A.

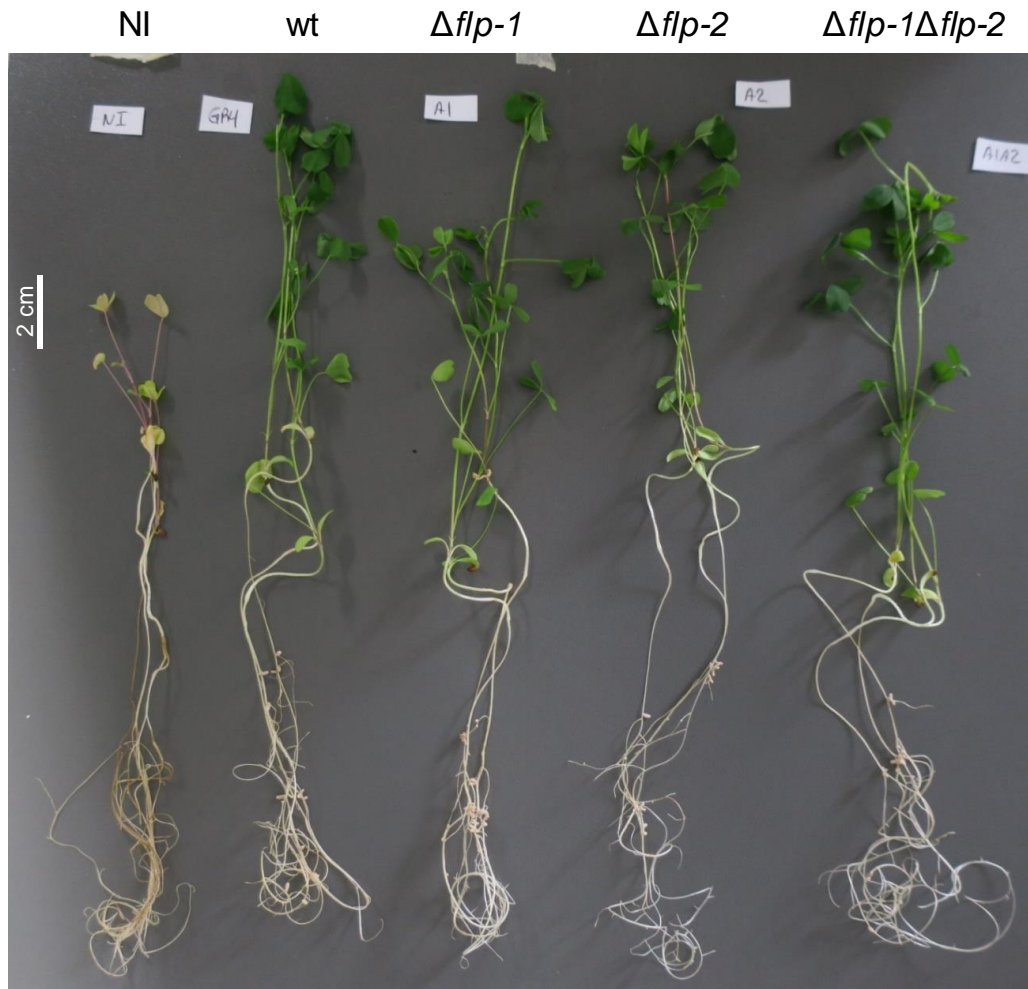


Figure S8. Appearance of alfalfa plants inoculated with *S. meliloti flp* mutant strains at the end of the nodulation kinetics experiment. Alfalfa plants were grown in hydroponic culture under axenic conditions in glass tubes. NI, non-inoculated; wt, wild-type strain GR4.

Table S1. Percent sequence identities (in parenthesis) among *S. meliloti* T4cP paralogues.

Locus GR4/Rm1021	Description	Chromosome <i>flp-1</i>	pSymA <i>flp-2</i>	pSymA Locus 3
Chromosome <i>flp-1</i>				
C770_GR4Chr0141/ <i>smc04118</i>	CpaK-like pseudopilin	<i>smc04117</i> (35% in 84 aa stretch)		
C770_GR4Chr0142/ <i>smc04117</i>	CpaJ-like pseudopilin	<i>smc04118</i> (35% in 84 aa stretch) <i>smc04110</i> (<i>cpaD1</i>) (25% in 125 aa stretch)	<i>sma1572</i> (40% in 66 aa stretch)	
C770_GR4Chr0143/ <i>smc04115</i>	CpaI-like putative pili related protein	<i>smc04111</i> (<i>cpaC1</i>) (21% in 103 aa stretch)	<i>sma1580</i> (36% in 97 aa stretch)	
C770_GR4Chr0144/ <i>smc04114</i> (<i>pilA1</i>)	Flp pilin subunit	<i>smc02447</i> (59)	<i>sma1570</i> (<i>pilA2</i>) (60)	
C770_GR4Chr0145/ <i>smc04113</i> (<i>cpaA1</i>)	Putative prepilin peptidase		<i>sma1578</i> (<i>cpaA2</i>) (37)	
C770_GR4Chr0146/ <i>smc04112</i> (<i>cpaB1</i>)	Component of the outer membrane secretin complex		<i>sma1576</i> (<i>cpaB2</i>) (36)	
C770_GR4Chr0147/ <i>smc04111</i> (<i>cpaC1</i>)	Component of the outer membrane secretin complex	<i>smc04115</i> (21% in 75 aa stretch)	<i>sma1580</i> (26% in 102 stretch)	<i>sma0163</i> (<i>pilQ2</i>) (30)
C770_GR4Chr0148/ <i>smc04110</i> (<i>cpaD1</i>)	Component of the outer membrane secretin complex	<i>smc04117</i> (25)		
C770_GR4Chr0149/ <i>smc04109</i> (<i>cpaE1</i>)	Putative docking protein of the secretion system		<i>sma1573</i> (<i>cpaE2</i>) (41)	
C770_GR4Chr0150/ <i>smc02820</i> (<i>cpaF1</i>)	Pilus assembly ATPase		<i>sma1568</i> (<i>cpaF2</i>) (72)	
C770_GR4Chr0151/ <i>smc02821</i>	CpaG-like putative inner membrane platform protein	<i>smc02822</i> (23)	<i>sma1564</i> (45)	
C770_GR4Chr0152/ <i>smc02822</i>	CpaH-like putative inner membrane platform protein	<i>smc02821</i> (22)	<i>sma1562</i> (49)	
C770_GR4Chr0153/ <i>smc04116</i> (TRm22)	Insertion sequence IS <i>Rm22</i>			
C770_GR4Chr0154/ <i>smc02824</i>	CpaO-like TPR containing protein			
Chromosomal orphan loci				
C770_GR4Chr2732/ <i>smc02446</i>	Flp pilin subunit	<i>smc04114</i> (<i>pilA1</i>) (59)	<i>sma1570</i> (<i>pilA2</i>) (64% in 25 aa stretch)	
C770_GR4Chr2733/ <i>smc02447</i>	Hypothetical protein			
C770_GR4Chr3249/ <i>smc03784</i>	CpaM-like polysaccharide deacetylase			
pSymA <i>flp-2</i>				
C770_GR4pC0512/ <i>sma1580</i>	CpaI-like putative pili related protein	<i>smc04115</i> (36% in 97 aa stretch) <i>smc04111</i> (<i>cpaC1</i>) (26% in 102 aa stretch)		<i>sma0163</i> (<i>pilQ2</i>) (26% in 98 aa stretch)
C770_GR4pC0513/ <i>sma1579</i>	Putative pili related protein			
C770_GR4pC0514/ <i>sma1578</i> (<i>cpaA2</i>)	Putative prepilin peptidase	<i>smc04113</i> (<i>cpaA1</i>) (38)		
C770_GR4pC0515/ <i>sma1576</i> (<i>cpaB2</i>)	Component of the outer membrane secretin complex	<i>smc04112</i> (<i>cpaB1</i>) (36%)		
C770_GR4pC0516/ <i>sma1573</i> (<i>cpaE2</i>)	Putative docking protein of the secretion system	<i>smc04109</i> (<i>cpaE1</i>) (41)		
C770_GR4pC0517/ <i>sma1572</i>	CpaJ-like pseudopilin protein	<i>smc04117</i> (40% in 66 aa stretch)		
C770_GR4pC0518/ <i>sma1570</i> (<i>pilA2</i>)	Flp pilin subunit	<i>smc04114</i> (<i>pilA1</i>) (60) <i>smc02447</i> (64% in 25 aa stretch)		
C770_GR4pC0519/ <i>sma1568</i> (<i>cpaF2</i>)	Pilus assembly ATPase	<i>smc02820</i> (<i>cpaF1</i>) (72)		
C770_GR4pC0520/ <i>sma1564</i>	CpaG-like putative inner membrane platform protein	<i>smc02821</i> (46)		
C770_GR4pC0521/ <i>sma1562</i>	CpaH-like putative inner membrane platform protein	<i>smc02822</i> (49)		
pSymA Locus 3				
C770_GR4pC1351/ <i>sma0163</i> (<i>pilQ2</i>)	RcpA-like putative component of the outer membrane secretin complex	<i>smc04111</i> (<i>cpaC1</i>) (30)	<i>sma1580</i> (26% in 98 aa stretch)	
C770_GR4pC1350/ <i>sma0164</i>	Hypothetical lipoprotein			
C770_GR4pC1349/ <i>sma0166</i>	CpaO/TadD-like TPR containing protein			

Table S2. Strains and plasmids used in this study.

Strain or plasmid	Relevant characteristics	Reference
<i>Sinorhizobium meliloti</i>		
GR4	Wild type, Nod ⁺ , Fix ⁺	[57]
GR4flaAB	GR4 <i>flaAflaB</i> ; Hyg ^r	[10]
GRflp1	GR4 Δ <i>flp-1</i>	This work
GRflp2	GR4 Δ <i>flp-2</i>	This work
GRflp1flp2	GRflp1 Δ <i>flp-2</i>	This work
GRflaflp1	GRflp1 <i>flaAflaB</i> ; Hyg ^r	This work
GRflaflp2	GRflp2 <i>flaAflaB</i> ; Hyg ^r	This work
GRflaflp1flp2	GRflp1flp2 <i>flaAflaB</i> ; Hyg ^r	This work
<i>Escherichia coli</i>		
DH5 α	<i>supE44</i> , <i>DlacU169</i> , <i>f80</i> , <i>lacZDM</i> , <i>5hsdR171</i> , <i>recA1</i> , <i>endA1</i> , <i>gyrA96</i> , <i>thi-1</i> , <i>relA1</i>	Bethesda Research Lab
S17-1	<i>thi</i> , <i>pro</i> , <i>recA</i> , <i>hsdR</i> , <i>hsdM</i> , <i>RP4-2-Tc::Mu-Km::Tn7</i>	[82]
Plasmids		
pCR-XL-TOPO	Cloning vector; Km ^r	Invitrogen
pTOPO- Δ flp1	pCR-XL-TOPO carrying the deleted version of the <i>flp-1</i> cluster; Km ^r	This work
pTOPO- Δ flp2	pCR-XL-TOPO carrying the deleted version of the <i>flp-2</i> cluster; Km ^r	This work
pK18 <i>mobsacB</i>	Suicide plasmid; Km ^r	[83]
pK18flaAB::Hyg	pK18 <i>mobsacB</i> derivative carrying the mutation <i>flaAflaB</i> ::Hyg; Km ^r , Hyg ^r	[10]
pK18- Δ flp1	pK18 <i>mobsacB</i> derivative carrying the deleted version of the <i>flp-1</i> cluster; Km ^r	This work
pK18- Δ flp2	pK18 <i>mobsacB</i> derivative carrying the deleted version of the <i>flp-2</i> cluster; Km ^r	This work
pGUS3	Plasmid containing <i>nfeD</i> :: <i>gusA</i> , [87] Km ^r	

Hyg^r and Km^r indicate hygromycin and kanamycin resistance, respectively.

Table S3. List of primers used in this study.

Name	Sequence (5'-3') ^a	Use
Rev	CAGGAAACAGCTATGAC	Sequencing
Univ	GTTTCCCAGTCACGAC	Sequencing
delpilA1.1	AAAGGATCCGTCTTAATTGCCGCTTGAGG (<i>Bam</i> HI)	Deletion in <i>flp-1</i>
delpilA1.2	GCTTGTCTCGTTGCGTCCTTCATCAGG	
delpilA1.3	CCTGATGAAGGACGCAACGAGGACAAGC	
delpilA1.4	AAAGGATCCCTCGAACGGACCATGACATCG (<i>Bam</i> HI)	
delpilA2.1	AAAAAGCTTCTGTTTCGGAGTCAGTCAGG (<i>Hind</i> III)	Deletion in <i>flp-2</i>
delpilA2.2	GTTCTCCTGTGCCAGCCGTATTCGATGG	
delpilA2.3	CCATCGAATACGGCTGGCACAGGAGAAC	
delpilA2.4	AAAAAGCTTTCCTCTCAACGATCAAAGC (<i>Hind</i> III)	

^aRestriction sites are underlined