Identifizierung neuer nicht-kodierender RNAs in *Xanthomonas campestris* pv. *vesicatoria* und funktionelle Charakterisierung der regulatorischen RNA sX13

Dissertation

zur Erlangung des
Doktorgrades der Naturwissenschaften (Dr. rer. nat.)

der

Naturwissenschaftlichen Fakultät I – Biowissenschaften –

der Martin-Luther-Universität
Halle-Wittenberg,

vorgelegt

von Herrn Cornelius Schmidtke

geb. am 15.05.1982 in Jena

Gutachter:

Prof. Dr. U. Bonas
Prof. Dr. G. Sawers
Prof. Dr. W. Hess

Verteidigung: 29.04.2014

Teile dieser Arbeit wurden in Fachzeitschriften publiziert:

Schmidtke, C., Findeiß, S., Sharma, C.M., Kuhfuss, J., Hoffmann, S., Vogel, J., Stadler, P.F. and Bonas, U. (2012) Genome-wide transcriptome analysis of the plant pathogen *Xanthomonas* identifies sRNAs with putative virulence functions. *Nucleic Acids Res.*, **40**, 2020-2031.

Findeiß, S., Schmidtke, C., Stadler, P.F. and Bonas, U. (2010) A novel family of plasmid-transferred anti-sense ncRNAs. *RNA Biol*, **7**, 120-124.

Schmidtke, C., Abendroth, U., Brock, J., Serrania, J., Becker, A. and Bonas, U. (2013) Small RNA sX13: a multifaceted regulator of virulence in the plant pathogen *Xanthomonas*. *PLoS Pathog.*, **9**, e1003626.

Zusammenfassung

Das Gram-negative pflanzenpathogene γ-Proteobakterium *Xanthomonas campestris* pv. *vesicatoria* ist der Erreger der bakteriellen Fleckenkrankheit auf Paprika und Tomate. Die Kenntnis bakterieller Faktoren, die zur Infektion von Wirtspflanzen beitragen, war zu Beginn dieser Arbeit auf Proteine begrenzt, wohingegen die Rolle nicht-kodierender RNAs in der Virulenz von *Xanthomonas* Spezies unbekannt war. Mittels eines cDNA-Sequenzieransatzes, welcher die Unterscheidung von Primärtranskripten und prozessierten RNAs ermöglicht, wurden 1.421 potentielle Transskriptionsstartpositionen sowie abundante nicht-kodierende RNAs in *X. campestris* pv. *vesicatoria* Stamm 85-10 identifiziert. Insgesamt wurden 24 potentiell regulatorische RNAs experimentell bestätigt, von denen drei (PtaRNA1, sX12 und sX13) näher untersucht wurden. Bioinformatische Analysen deuten darauf hin, dass der *ptaRNA1* ('plasmid-transferred antisense RNA 1') Lokus durch horizontalen Gentransfer verbreitet wird und lassen vermuten, dass die PtaRNA1 antisense RNA die Synthese eines potentiell toxischen Proteins unterdrückt. Acht der in dieser Arbeit identifizierten nicht-kodierenden RNAs, einschließlich sX12, sind mit dem Typ III Sekretionssystem, einem essentiellen Pathogenitätsfaktor von *X. campestris* pv. *vesicatoria*, ko-reguliert. Durch genetische Analysen konnte nachgewiesen werden, dass sX12 die Virulenz von *X. campestris* pv. *vesicatoria* fördert. In dieser Arbeit wurde zudem die konstitutiv exprimierte und abundante sX13 RNA als neuartiger Virulenzfaktor von *X. campestris* pv. *vesicatoria* identifiziert. sX13 fördert die Expression von Komponenten und Substraten des Typ III Sekretionssystems und trägt zum bakteriellen Wachstum in Kultur bei. 'Microarray' Analysen ergaben ein großes sX13 Regulon und lassen vermuten, dass sX13 zur Adaption von *X. campestris* pv. *vesicatoria* an sich verändernde Umweltbedingungen beiträgt. sX13 hemmt die Expression des RNA-Bindeproteins Hfq, welches in zahlreichen Bakterien für die Aktivität regulatorischer RNAs essentiell ist und zur Virulenz pathogener Bakterien beiträgt. Die Ergebnisse deuten darauf hin, dass sX13 Hfq-unabhängig agiert und dass *hfq* für die Virulenz von *X. campestris* pv. *vesicatoria* entbehrlich ist. Strukturanalysen von sX13 sowie Deletions- und Komplementationsexperimente ergaben, dass sX13 drei 'Stem-Loop' Strukturen mit 'C'-reichen Loops aufweist, welche in unterschiedlichem Maße zur Virulenz von *X. campestris* pv. *vesicatoria* beitragen. Mittels eines GFP-Reportersystems wurde nachgewiesen, dass 'C'-reiche sX13 Loops und 'G'-reiche Motive in potentiellen Ziel-mRNAs für die sX13-abhängige Repression der Proteinsynthese essentiell sind.

Summary

The Gram-negative plant-pathogenic γ-proteobacterium *Xanthomonas campestris* pv. *vesicatoria* is the causal agent of bacterial spot disease on pepper and tomato. At the beginning of this study, the knowledge of bacterial factors, which contribute to the infection of host plants, was limited to proteins, whereas the role of noncoding RNAs in the virulence of *Xanthomonas* species was unknown. Using a cDNA-sequencing approach, which allows distinguishing unprocessed and processed RNAs, 1,421 putative transcription start sites and abundant noncoding RNAs were identified in *X. campestris* pv. *vesicatoria* strain 85-10. In total, 24 putative regulatory RNAs were experimentally verified, three of which (PtaRNA1, sX12 and sX13) were analyzed in more detail. Bioinformatic analyses suggest that the *ptaRNA1* ('plasmid-transferred antisense RNA 1') locus is transferred via horizontal gene transfer and further indicate that the PtaRNA1 antisense RNA represses the synthesis of a presumably toxic protein. Eight of the identified noncoding RNAs, including sX12, are co-regulated with the type III secretion system, which constitutes an essential pathogenicity factor of *X. campestris* pv. *vesicatoria*. Genetic analyses showed that sX12 contributes to virulence of *X. campestris* pv. *vesicatoria*. Furthermore, this work revealed that the constitutively expressed and abundant sX13 RNA represents a novel virulence factor of *X. campestris* pv. *vesicatoria*. sX13 promotes the expression of components and substrates of the type III secretion system and contributes to bacterial growth in culture. Microarray analyses revealed a large sX13 regulon and suggest that sX13 contributes to environmental adaptation of *X. campestris* pv. *vesicatoria*. sX13 inhibits the expression of the RNA-binding protein Hfq, which is essential for the activity of regulatory RNAs in many bacteria and contributes to virulence of pathogenic bacteria. The data suggest that sX13 acts Hfq-independently. Furthermore, *hfq* is presumably not involved in virulence of *X. campestris* pv. *vesicatoria*. Structure analyses of sX13 and deletion and complementation experiments revealed that sX13 consists of three stem-loops with 'C'-rich loops, which differentially contribute to virulence of *X. campestris* pv. *vesicatoria*. Using a GFP-reporter system, both the 'C'-rich sX13 loops and 'G'-rich motifs in presumed target mRNAs were shown to be essential for the sX13-dependent repression of protein synthesis.

Danksagung

Mein besonderer Dank gebührt allen, die an mich geglaubt haben, allen voran Frau Prof. Dr. Ulla Bonas für die Bereitstellung dieses hochinteressanten und anspruchsvollen Forschungsthemas, die fruchtbaren Diskussionen und ihr stetiges Vertrauen in meine Fähigkeiten.

Zudem danke ich allen Kooperationspartnern, insbesondere Sven Findeiß, Juliane Brock und Ulrike Abendroth, die maßgeblich zum Erfolg dieser Arbeit beigetragen haben.

Für die schöne Zeit, den regen Gedankenaustausch und die freundliche Arbeitsatmosphäre bedanke ich mich bei den Mitgliedern des Labors 215, Evelyn Löschner, Ulrike Abendroth, Juliane Brock, Christine Wagner und Johannes Stuttmann, sowie bei allen Mitgliedern der Arbeitsgruppe Bonas. Carola Kretschmer, Hannelore Espenhahn und Marina Schulze danke ich für die hervorragende technische Assistenz und Bianca Rosinsky für ihren grünen Daumen.

Ein herzlicher Dank gilt Heike Berndt, Daniela Büttner, Steve Schulz, Sebastian Schulze, Tom Schreiber und Oliver Müller für die pausenfüllenden Diskussionen und Hilfe in allen Lebenslagen. Simone Hahn und Robert Szczesny danke ich für wahre Freundschaft.

Mein aufrichtiger Dank gilt meiner Familie für ihre bedingungslose Unterstützung und meiner Frau Katja, ohne deren Liebe, Geduld und Zuspruch diese Arbeit vermutlich nicht möglich gewesen wäre. Danke!

Inhaltsverzeichnis

Abbildungsverzeichnis

Abkürzungsverzeichnis

Abb.	Abbildung
Ado-Cbl	Adenosylcobalamin
AS	Aminosäure
asRNA	*cis*-kodierte antisense RNA
avr, Avr	Avirulenz
cDNA	‚complementary DNA‘
di-GMP	di-Guanosinmonophosphat
DNA	‚deoxyribonucleic acid’, Desoxyribonukleinsäure
dRNA-Seq	‚differential RNA sequencing‘
DSF	‚diffusible signal factor’
ECW	‚Early Californian Wonder‘, Kultivar von *Capsicum annuum*
fMET-tRNA$^{\text{fMet}}$	N-formyl-methionyl tRNA
FMN	Flavinmononucleotid
gfp, GFP	‚green fluorescent protein‘
hpa, Hpa	‚*hrp*-associated‘
HR	hypersensitive Reaktion
hrc, Hrc	‚*hrp*-conserved‘
hrp, Hrp	‚hypersensitive response and pathogenicity‘
Kb, Kbp	Kilobasen, Kilobasenpaare
lmRNA	‚leaderless mRNA’
Mbp	Megabasenpaare
mRNA	‚messenger RNA‘
ncRNA	‚non-coding RNA‘, nicht-kodierende RNA
NGS	‚next-generation sequencing‘
Nt	Nukleotide
NYG	‚nutrient-yeast-glycerol‘
ORF	‚open reading frame‘, offenes Leseraster
PIP	‚plant inducible promoter‘
pRNA	‚product RNA‘
ptaRNA1, PtaRNA1	‚plasmid-transferred antisense RNA 1‘
pv.	Pathovar
qRT-PCR	quantitative ‚reverse transcription‘-PCR
QS	‚quorum sensing‘
RBS	Ribosomenbindestelle
RNA	‚ribonucleic acid’, Ribonukleinsäure

RNase	Ribonuklease
RNA-Seq	‚RNA sequencing‘
rpf, Rpf	‚regulation of pathogenicity factors‘
rRNA	ribosomale RNA
SAH	*S*-Adenosylhomocystein
SAM	*S*-Adenosylmethionin
SD	Shine-Dalgarno
spp.	*species pluralis*, Spezies
sRNA	‚small RNA‘, *trans*-kodierte RNA
subsp.	Subspezies
T3S System	Typ III Sekretionssystem
TA	Toxin-Antitoxin
TEX	Terminator-Exonuklease
Tfp	Typ IV Pilus
TLS	Translationsstartcodon
tmRNA	‚transfer-messenger RNA‘
TPP	Thiaminpyrophosphat
tRNA	‚transfer RNA’
TSS	‚transcription start site‘,Transkriptionsstartposition
UTR	untranslatierte Region
Xac	*Xanthomonas axonopodis* pv. *citri*
Xal	*Xanthomonas albilineans*
Xam	*Xanthomonas axonopodis* pv. *manihotis*
Xcc	*Xanthomonas campestris* pv. *campestris*
Xcv	*Xanthomonas campestris* pv. *vesicatoria*
Xoc	*Xanthomonas oryzae* pv. *oryzicola*
Xoo	*Xanthomonas oryzae* pv. *oryzae*
xop, Xop	‚*Xanthomonas* outer protein‘

1. Einleitung

Bakterien sind extrem anpassungsfähige Organismen, die nahezu alle denkbaren Lebensräume besiedeln. Eine Voraussetzung hierfür ist die Fähigkeit, Umweltsignale zu perzipieren und auf Veränderungen zu reagieren. Neben der Reaktion auf abiotische Faktoren wie Temperatur, Osmolarität und Nährstoffverfügbarkeit müssen pathogene (krankheitserregende) Bakterien in der Lage sein, sich der Immunabwehr eukaryotischer Wirte zu entziehen bzw. diese zu unterdrücken. Die bakterielle Anpassung wird durch transkriptionelle Regulation der Genexpression ermöglicht sowie durch regulatorische RNAs, welche die Genexpression auf posttranskriptioneller Ebene modulieren. Eine regulatorische Funktion von RNA Molekülen wurde bereits vor 50 Jahren von Jacob und Monod vermutet (89). Erst zwei Jahrzehnte später wurden in *Escherichia coli* RNAs identifiziert, welche die Plasmidreplikation bzw. die Aktivität von Transposons unterdrücken (201,211,219). Heute ist bekannt, dass bakterielle RNA-Regulatoren zahlreiche physiologische Prozesse modulieren, z.B. die Reaktion auf verschiedene Stressbedingungen und die Aufnahme und Verwertung von Kohlenstoffquellen (67,174). In pathogenen Bakterien beeinflussen regulatorische RNAs zudem die Expression von Genen, welche die Infektion von Wirtsorganismen fördern (Virulenzgene)(160). Bakterielle regulatorische RNAs sind überwiegend kurze Transkripte (50-300 Nukleotide; Nt), die meist kein Protein kodieren (,non-coding RNA'; ncRNA) und aufgrund intramolekularer Basenpaarungen stabile Sekundärstukturen ausbilden. Durch komplementäre Basenpaarung mit proteinkodierenden RNAs (,messenger RNA'; mRNA), Interaktion mit Proteinen oder Vermittlung von Strukturveränderungen in mRNAs beeinflussen ncRNAs Prozesse wie Transkription, Translation sowie die Stabilität von Transkripten (210,245). Im Folgenden werden Ansätze zur Identifizierung bakterieller ncRNAs sowie deren Funktionsweisen näher betrachtet.

1.1. Experimentelle Identifizierung regulatorischer RNAs

Die ersten regulatorischen RNAs, z.B. Spot42 und MicF, wurden durch Zufall in *E. coli* entdeckt (139,183). Mit zunehmender Verfügbarkeit bakterieller Genomsequenzen wurden ncRNAs durch vergleichende Genomanalysen vorhergesagt (164). Die Mehrheit der heute bekannten ncRNAs wurde mittels ,Microarray' Analysen oder durch Sequenzierung von cDNA (,complementary DNA') identifiziert (198). ,Microarray' Analysen basieren auf der Hybridisierung von immobilisierten Oligonukleotiden mit fluoreszenzmarkierter cDNA. Die Oligonukleotide sind hierbei gegen mRNAs oder ncRNAs gerichtet (243) oder decken systematisch den ,sense' und ,antisense' Strang eines Genoms ab (,tiling arrays'). Mittels ,tiling arrays' wurden beispielsweise im Humanpathogen *Listeria monocytogenes* ncRNAs identifiziert, die unter Infektionsbedingungen von Bedeutung sind (218). Sogenannte ,RNomics'-Ansätze zur Sequenzierung von klonierter cDNA (100,234) wurden mittlerweile von Hochdurchsatz-Sequenzierungen abgelöst. Diese modernen Sequenziertechniken

werden als NGS ('next-generation sequencing') oder RNA-Seq ('RNA sequencing') bezeichnet (101). Durch Ligation von RNA-Adaptorsequenzen an die 5'- und 3'-Enden von Transkripten und cDNA-Synthese mittels 3'-adaptorspezifischen Oligonukleotiden erlauben RNA-Seq Analysen die strangspezifische Zuordnung von Sequenzierdaten ('reads') zu genomischen Sequenzen ('mapping'). Neben Sequenziertechnologien wie SOLiD (199) und IonTorrent (181)(http://www.lifetechnologies.com) wurde überwiegend 454-Pyrosequenzierung (Roche; www.454.com)(131) und Illumina/ Solexa Sequenzierung (www.illumina.com)(10) für bakterielle Transkriptomstudien genutzt (166).

Bis heute wurden durch RNA-Seq Analysen hunderte von bakteriellen ncRNA Kandidaten identifiziert, z.B. in den Gram-negativen Humanpathogenen *Salmonella enterica* und *Legionella pneumophila* sowie in den Gram-positiven Pathogenen *L. monocytogenes* und *Staphylococcus aureus* (14,146,202,251). ncRNA Kandidaten werden üblicherweise mittels Northern Blot Analysen validiert.

1.2. Identifizierung von Transkriptionsstarts mittels dRNA-Seq

Bakterielle Primärtranskripte weisen eine Triphosphatgruppe am 5'-Ende auf (5'-PPP), wohingegen prozessierte oder degradierte RNA durch ein 5'-Monophosphat (5'-P) oder eine 5'-Hydroxylgruppe gekennzeichnet ist. Die Unterscheidung dieser RNA Spezies ist essentiell für die Identifizierung von Transkriptionsstartpositionen ('transcription start site', TSS), da auch stabile RNA Prozessierungsprodukte, wie ribosomale RNA (rRNA), einheitliche 5'-Enden aufweisen. Im Gegensatz zu RNA-Seq Ansätzen erlaubt die sogenannte dRNA-Seq Methode ('differential RNA sequencing') die Identifizierung von Primärtranskripten (196). Hierbei werden mittels einer 5'-P-abhängigen Terminator-Exonuklease (TEX) zunächst Transkripte degradiert, die eine 5'-P-Gruppe tragen; Primärtranskripte mit 5'-PPP-Gruppen werden dadurch angereichert. Durch anschließende Behandlung mit TAP ('tobacco acid pyrophosphatase') werden 5'-PPP- in 5'-P-Gruppen konvertiert und dann mit Adaptorsequenzen ligiert. Nach der cDNA-Sequenzierung können TSSs anhand der Anreicherung von 'reads' in der mittels TEX-Behandlung generierten cDNA Bibliothek gegenüber der unbehandelten cDNA Bibliothek identifiziert werden. Diese Methode wurde erstmals für das Humanpathogen *Helicobacter pylori* angewendet und identifizierte TSSs für 87,5% der annotierten proteinkodierenden Gene sowie mehr als 60 ncRNAs (196).

1.3. Funktionen und Mechanismen RNA-vermittelter Regulation

Bakterielle regulatorische RNAs werden anhand ihrer Funktionsweise in vier Hauptkategorien unterteilt: (i) Riboswitches und RNA-Thermometer, (ii) proteinbindende RNAs, (iii) *cis*-kodierte antisense RNAs sowie (iv) *trans*-kodierte RNAs. Darüber hinaus verfügen zahlreiche Bakterien über ein RNA-basiertes adaptives Immunsystem: Sogenannte CRISPR (,clustered regularly interspaced short palindromic repeats')-Elemente schützen Bakterienzellen vor eindringenden Viren und Plasmiden (205) und werden im Folgenden nicht näher betrachtet.

1.3.1. Riboswitches und RNA-Thermometer

Riboswitches sind *cis*-regulatorische RNA Elemente in der 5'-untranslatierten Region (UTR) von meist polycistronischen mRNAs (Abb. 1A)(245). Riboswitches nehmen durch Bindung spezifischer Metabolite Einfluss auf die Transkription oder Translation der stromabwärts lokalisierten Gene, welche meist an der Biosynthese oder dem Transport des jeweiligen Liganden beteiligt sind. Verschiedene Klassen von Riboswitches binden Koenzyme oder deren Derviate, z.B. Thiaminpyrophosphat (TPP), Flavinmononucleotid (FMN), *S*-Adenosylmethionin (SAM), *S*-Adenosylhomocystein (SAH) oder Adenosylcobalamin (Ado-Cbl)(5,79). Andere Riboswitches binden Aminosäuren wie Glycin oder Lysin bzw. Nucleobasen wie Adenin oder Guanin (5,79). Zudem wurden Riboswitches identifiziert, die den sekundären Botenstoff zyklisches di-Guanosinmonophosphat (zyklisches di-GMP) bzw. Mg^{2+} binden (79).

Die Bindung eines Liganden an die evolutionär konservierte Aptamerregion (35-200 Nt) eines Riboswitches induziert eine strukturelle Veränderung in der sogenannten Expressionsplattform. Dies beeinflusst entweder die weitere Transkription der mRNA durch Ausbildung oder Auflösung einer transkriptionellen Terminatorstruktur im 5'-UTR oder die Zugänglichkeit der Ribosomenbindestelle (RBS) für die Ausbildung des Translations-Initiationskomplexes (Abb. 1A)(193). In Bakterien beginnt die Translation mit der Bindung eines Komplexes aus 30S Ribosomenuntereinheiten, der Initiator transfer RNA (N-formyl-methionyl tRNA, fMET-tRNA[fMet]) und Initiationsfaktoren an die Shine-Dalgarno (SD) Sequenz der mRNA (70,111). Die SD-Sequenz (Konsensus in *E. coli* ,GGAGG') ist Teil der RBS und lokalisiert wenige Nukleotide stromaufwärts des Translationsstartcodons (TLS)(200). Durch komplementäre Basenpaarung mit dem 3'-Ende der 16S rRNA, der sogenannten Anti-SD Sequenz (,CCUCC'), rekrutiert die SD-Sequenz 30S Ribosomenuntereinheiten an die mRNA (86,90,175,208). Nach Ausbildung des Initiationskomplexes binden 50S Ribosomenuntereinheiten, wodurch translationsaktive 70S Ribosomenkomplexe ausgebildet werden.

RNA-Thermometer sind regulatorische Elemente, die in 5'-UTRs von temperaturresponsiven Genen lokalisiert sind und temperaturabhängig ihre Faltung verändern (107). Niedrige Temperaturen bedingen meist eine Konformation, welche die Bindung von 30S-Ribosomenuntereinheiten an die

RBS der mRNA verhindert. Dagegen vermitteln höhere Temperaturen das Aufschmelzen inhibitorischer Sekundärstrukturen (107). Sogenannte ‚ROSE'-Elemente kontrollieren die Synthese von Hitzeschockproteinen (149), während ‚FourU'-Elemente u.a. die Expression von Virulenzgenen regulieren (238). Die temperaturabhängige Expression von Virulenzgenen, z.B. *prfA* in *L. monocytogenes* und *lcrF* in *Yersinia pseudotuberculosis*, ermöglicht pathogenen Bakterien die Erkennung und Infektion warmblütiger Wirtsorganismen (12,93).
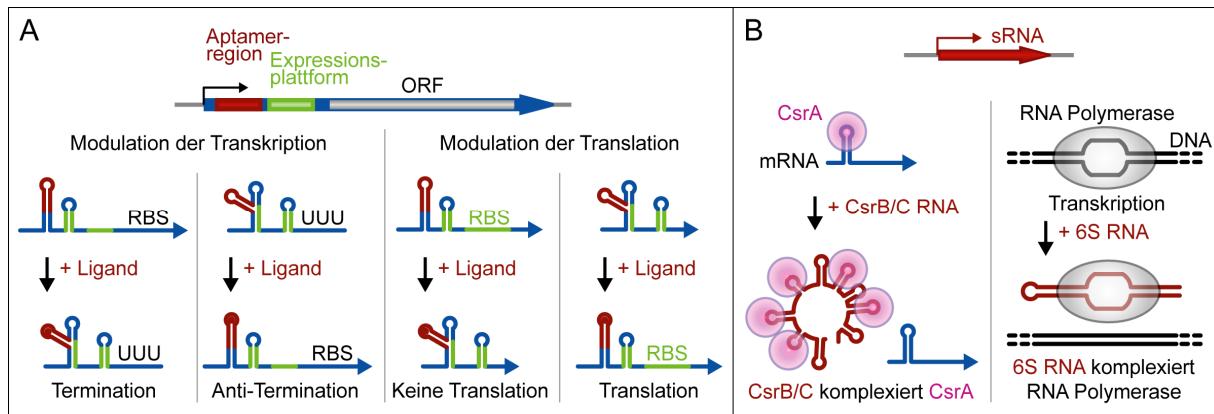
Abbildung 1. Modelle der Funktionsweise von Riboswitches und proteinbindenden RNAs.
(A) Riboswitches sind in 5'-UTRs von mRNAs (blau) lokalisiert und umfassen eine Liganden-bindende Aptamerregion sowie eine Expressionsplattform. (Linke Seite) Die Bindung des Liganden fördert oder hemmt die Ausbildung einer transkriptionellen Terminatorstruktur (UUU) im 5'-UTR. (Rechte Seite) Die Ligandenbindung fördert die Ausbildung oder Auflösung einer Sekundärstruktur, welche die Ribosomenbindestelle (RBS) blockiert und die Translation des offenen Leserasters (ORF) hemmt. (B) Modulation der Proteinaktivität durch ncRNAs. (Linke Seite) Die Aktivität des CsrA Proteins, welches die Translation von Ziel-mRNAs beeinflusst, wird durch Bindung der ncRNA CsrB bzw. CsrC gehemmt. (Rechte Seite) Die δ^{70}-assoziierte RNA Polymerase ermöglicht die Transkription durch Bindung an Promotorregionen. Die DNA-Assoziation der RNA Polymerase wird durch Bindung der 6S RNA gehemmt und führt zur verminderten Aktivität von δ^{70}-Promotoren. (Abb. modifiziert nach Waters und Storz, 2009 (245)).

1.3.2. RNA-vermittelte Modulation der Proteinaktivität

Regulatorische RNAs können essentielle Funktionen von Ribonukleoproteinkomplexen vermitteln (z.B. rRNA und ‚transfer-messenger RNA', tmRNA) oder die Aktivität gebundener Proteine modulieren (z.B. 6S RNA und CsrB/ CsrC)(Abb. 1B). Das namensgebende Charakteristikum der tmRNA ist eine tRNA- sowie eine mRNA-ähnliche Region, welche ein kurzes offenes Leseraster (‚open reading frame', ORF) enthält (142). Bei Unterbrechung des Translationsprozesses bindet tmRNA an Ribosomen und terminiert die Translation. Hierbei vermittelt das tmRNA-kodierte Polypeptid den Abbau der unvollständig translatierten Polypeptidkette, während das Stoppcodon des tmRNA-ORFs die Ablösung des Ribosoms von der mRNA ermöglicht (142).

Die hochkonservierte 6S RNA (180-200 Nt) weist eine Struktur auf, die der Konformation der DNA während der Transkription ähnelt (Abb. 1B)(222). In *E. coli* akkumuliert die 6S RNA in der stationären Wachstumsphase und bindet die mit dem Sigmafaktor δ^{70} assoziierte RNA-Polymerase

(244). Infolgedessen werden Gene mit δ^{70} Promotoren vermindert transkribiert (221). 6S RNA kann auch als Matrize der RNA-Polymerase dienen und generiert 14-20-Nt ‚product RNAs' (pRNAs). Die pRNA Transkription wird vermutlich durch einen Anstieg der Nukleosidtriphosphat-Konzentration induziert und vermittelt die Ablösung der 6S RNA von der RNA-Polymerase (242).

Die zentrale Rolle von ncRNAs in der Regulation physiologischer Prozesse wird insbesondere anhand des Csr- (‚carbon storage regulator') bzw. des verwandten Rsm (‚repressor of secondary metabolites')-Systems deutlich (Abb. 1B). Das RNA-Bindeprotein CsrA reguliert in *E. coli* die Synthese und Verwertung von Kohlenstoffquellen sowie die Motilität (177,248). CsrA inhibiert die Translation der meisten Ziel-mRNAs oder beeinflusst deren Stabilität durch Bindung an multiple ‚GGA'-Sequenzmotive in den 5'-UTRs (3,217). Die CsrA Aktivität wird durch die ncRNAs CsrB und CsrC moduliert, welche multiple ‚GGA'-Motive enthalten und mit mRNAs um die Bindung an CsrA konkurrieren (Abb. 1 B)(3,120,250). Orthologe des *E. coli* Csr-Systems wurden beispielsweise in *Pseudomonas*-, *Legionella*- und pflanzenpathogenen *Erwinia*- und *Xanthomonas* spp. identifiziert und sind u.a. an der Regulation der Zelldichte-abhängigen Genexpression, der Motilität und der Virulenzgenexpression beteiligt (34,110,264).

1.3.3. *Cis*-kodierte antisense RNAs

Cis-kodierte antisense RNAs (asRNAs) werden vom DNA-Gegenstrang proteinkodierender Gene transkribiert und weisen daher perfekte Komplementarität, meist über mehr als 75 Nt, zur korrespondierenden mRNA auf (Abb. 2A)(20). RNA-Seq Analysen ergaben, dass Bakterien eine unerwartet hohe Zahl von asRNAs exprimieren (64), z.B. wurden in *H. pylori* asRNAs für 46% der annotierten ORFs identifiziert (196). Die Transkriptlängen von asRNAs variieren und reichen von etwa 100 Nt, wie SymR und GadY in *E. coli* (99,154), bis mehrere Kilobasen (Kb), z.B. asRNAs im Cyanobakterium *Prochlorococcus* sp. Stamm MED4 (207). asRNA Gene können mit dem 5'- oder 3'-UTR oder dem ORF der *cis*-lokalisierten Gene überlappen, wobei die Interaktion von asRNA und mRNA in einer veränderten Translation der mRNA bzw. einer veränderten Stabilität der Transkripte resultiert (Abb. 2A)(20,64). Die *E. coli* asRNAs SymR und GadY gehören zu den am besten untersuchten asRNAs. SymR überlappt in antisense Orientierung mit der RBS und dem TLS der *symE* mRNA und unterdrückt deren Translation (99). GadY vermittelt die Prozessierung der bicistronischen *gadXW* mRNA zwischen *gadX* und *gadW*, wobei die prozessierten Transkripte eine höhere Stabilität als die unprozessierte mRNA aufweisen (154,220). Die asRNA-induzierte mRNA Prozessierung wird meist durch die Ribonuklease (RNase) E oder RNase III vermittelt (64). RNase III spaltet präferentiell perfekt gepaarte RNA-RNA Komplexe, wohingegen RNase E bevorzugt imperfekt gepaarte Komplexe degradiert. Neben der Modulation der Translation und mRNA Stabilität können asRNAs auch die Transkription beeinflussen (Abb. 2A). Beispielsweise vermittelt im Fischpathogen *Vibrio anguillarum* die asRNA RNAß die vorzeitige Termination der Transkription des *fatDCBA-angRT*

Operons stromabwärts von *fatA* (209). Die regulatorische Funktion einiger anderer asRNAs beruht vermutlich allein auf deren Transkription, da divergent transkribierte Promotoren einander beeinflussen können (transkriptionelle Interferenz)(64). Die Transkription eines DNA-Strangs durch die RNA Polymerase verhindert hierbei die Initiation bzw. Elongation der Transkription auf dem DNA-Gegenstrang (156).
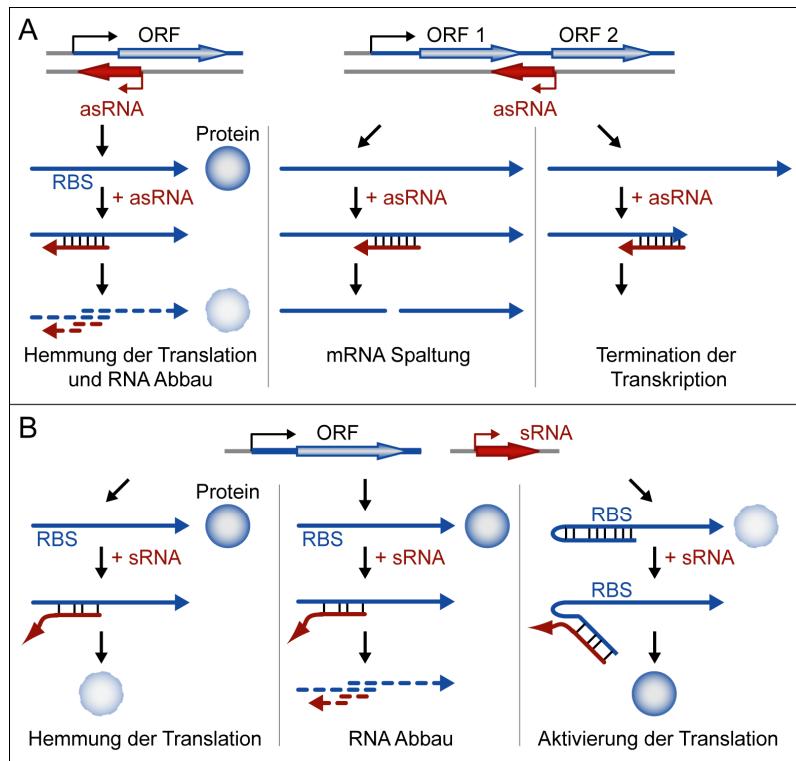
Abbildung 2. Regulatorische Mechanismen basenpaarender RNAs.
(A) *cis*-kodierte asRNAs. (Linke Seite) Die Interaktion einer asRNA mit der Ribosomenbindestelle (RBS) der Ziel-mRNA hemmt die Initiation der Translation und induziert meist den RNA-Abbau. (Mitte) asRNAs können die RNase-vermittelte Prozessierung polycistronischer mRNAs induzieren, wobei die prozessierten Transkripte eine veränderte Stabilität aufweisen. (Rechte Seite) Während der Transkription der polycistronischen mRNA kann die Bindung einer asRNA die vorzeitige Termination der Transkription vermitteln und die Expression stromabwärts lokalisierter Cistrons unterdrücken. (B) *trans*-kodierte sRNAs. sRNAs interagieren meist über kurze und imperfekt-komplementäre Sequenzen mit den 5'-UTRs von Ziel-mRNAs und hemmen deren Translation (linke Seite), induzieren den RNase-vermittelten Abbau der mRNA (Mitte) oder fördern die Translation durch Auflösung inhibitorischer Sekundärstrukturen (rechte Seite). (Abb. modifiziert nach Waters und Storz, 2009 (245)).

1.3.4. *Trans*-kodierte RNAs

Die Mehrheit der charakterisierten ncRNAs wurde in den Enterobakterien *E. coli* und *Salmonella* untersucht und geht Basenpaarungen mit mRNAs ein, welche im Genom abseits der ncRNA kodiert sind (in *trans*). Diese Transkripte werden als *trans*-kodierte RNAs oder ‚small RNAs' (sRNAs) bezeichnet, sind überwiegend nicht-kodierend und stark strukturiert und weisen Längen von meist 50-300 Nt auf (210). Im Gegensatz zu *cis*-kodierten asRNAs interagieren sRNAs üblicherweise über kurze und imperfekt komplementäre Sequenzen (~10-25 Nt) mit den 5'-UTRs von multiplen mRNAs (67,245). Dies erschwert die bioinformatische Vorhersage von Ziel-mRNAs (117). Da die meisten sRNAs die Translation und/ oder die Stabilität von mRNAs modulieren, können potentielle Ziel-mRNAs durch Proteom- und Transkriptomanalysen von sRNA-Deletionsmutanten oder Überexpressionsstämmen identifiziert werden (198,236). Der Einfluss von sRNAs auf Ziel-mRNAs wird häufig mittels translationaler mRNA-Reporterfusionen analysiert, z.B. Fusionen mit *gfp* (‚green fluorescent protein')(224).

1.3.4.1. Mechanismen sRNA-vermittelter Regulation

Die meisten sRNAs wirken negativ auf Ziel-mRNAs und inhibieren die Initiation der Translation durch Basenpaarung mit oder nahe der RBS der mRNA (Abb. 2B)(67). Zudem sind Beispiele bekannt, in denen sRNAs die Translation von Ziel-mRNAs unterdrücken, indem sie an Sequenzen binden, die bis zu 70 Nt stromaufwärts und 15 Nt stromabwärts des TLS lokalisiert sind (19,84). Parallel zur Hemmung der Translation induzieren sRNA-mRNA Interaktionen häufig den RNase E-vermittelten Abbau der beteiligten Transkripte (Abb. 2B)(31). Für die *E. coli* sRNAs RyhB und SgrS wurde nachgewiesen, dass die Hemmung der Ziel-mRNA Translation unabhängig von der RNase E-vermittelten Degradation erfolgt (145). Einige sRNAs fördern ausschließlich den Abbau von Ziel-mRNAs, z.B. bindet die *Salmonella* sRNA MicC in der kodierenden Region der *ompD* mRNA und beschleunigt deren Abbau durch RNase E (163). Neben RNase E wurde eine Rolle von RNase III in der Degradation von sRNA-mRNA Komplexen beschrieben (31,87,230).

sRNAs können auch als Aktivatoren der Genexpression wirken und die Translation von Ziel-mRNAs fördern (60). Hierbei induziert die Bindung einer sRNA an den 5'-UTR der mRNA die Auflösung einer inhibitorischen Sekundärstruktur, welche die RBS einschließt und die Initiation der Translation verhindert (Abb. 2B). Beispielsweise sind drei *E. coli* sRNAs bekannt (ArcZ, DsrA und RprA), welche die Synthese des Sigmafaktors RpoS aktivieren (125,126,203).

1.3.4.2. Das RNA-Chaperon Hfq

Hfq (‚host factor for the replication of the RNA phage Qß‘) wurde als *E. coli* Wirtsfaktor für die Replikation des Bakteriophagen Qß identifiziert (59). Das RNA-Bindeprotein Hfq ist in etwa 50% aller Bakterien konserviert und vermittelt, insbesondere in Enterobakterien, die Interaktion von sRNAs und Ziel-mRNAs (44,235). Für *Salmonella* wurde beschrieben, dass Hfq mit rund 100 sRNAs und etwa 20% der transkribierten mRNAs assoziiert ist (160,202). Hfq bildet eine homohexamere Ringstruktur aus und bindet an ‚AU‘-reiche einzelsträngige Sequenzen in sRNAs und mRNAs (235). Es wird vermutet, dass Hfq die lokale Konzentration von sRNAs und Ziel-mRNAs erhöht und dadurch die Basenpaarung von kurzen und imperfekt-komplementären Sequenzen begünstigt (44). Dafür spricht, dass sRNA-mRNA Interaktionen über längere und perfekt komplementäre Regionen, wie im Fall *cis*-kodierter asRNAs, in der Regel Hfq-unabhängig sind. Hfq ist mit RNase E und weiteren Proteinen des sogenannten RNA-Degradosoms assoziiert und vermittelt dadurch den Abbau von sRNA-mRNA Komplexen (32,88,144). Interessanterweise nimmt die Stabilität von *E. coli* sRNAs in Abwesenheit von Hfq ab, was vermuten lässt, dass Hfq-gebundene sRNAs vor Degradation geschützt sind (235).

Die Inaktivierung des *hfq* Gens ist in verschiedenen Bakterien mit pleiotropen Phänotypen, wie reduziertem Wachstum, veränderter Motilität sowie veränderter Toleranz gegenüber Stressbedingungen verbunden (35). Zudem beeinträchtigt die Inaktivierung von *hfq* die Virulenz zahlreicher humanpathogener Bakterien (35,160). Dagegen ist Hfq für die Virulenz von beispielsweise *L. pneumophila* und *S. aureus* entbehrlich, wenngleich sRNAs in diesen Bakterien die Virulenzgenexpression modulieren (13,133).

1.4. Gram-negative pflanzenpathogene Bakterien

Der Befall von Nutzpflanzen mit Schädlingen und Parasiten verursacht weltweit erhebliche Ertragsverluste und stellt ein ernstzunehmendes Problem für die Nahrungsmittelproduktion dar. Gram-negative pflanzenpathogene Bakterien sind vor allem in feucht-warmen aber auch in gemäßigten Klimaregionen von Bedeutung (130). Von besonderem ökonomischen und wissenschaftlichen Interesse sind *Ralstonia solanacearum*, der Erreger der bakteriellen Welke in mehr als 200 Pflanzenarten (63), *Erwinia amylovora*, der Erreger des Feuerbrands (228), das Tumor-induzierende Bakterium *Agrobacterium tumefaciens* (167) sowie Pathovare (pv.) von *Pseudomonas syringae*, welche Blattflecken, Brände oder Geschwüre auslösen (130). Ernteverluste durch Vertreter der Gattung *Xanthomonas* sind insbesondere schwerwiegend, da Wirtspflanzen wie Reis und Maniok die Nahrungsgrundlage von Millionen von Menschen darstellen (130). Im Folgenden werden Pathogenitätsmechanismen von pflanzenpathogenen Bakterien der Gattung *Xanthomonas* näher betrachtet.

1.4.1. Die Gattung *Xanthomonas*

Gram-negative γ-Proteobakterien der Gattung *Xanthomonas* sind stäbchenförmige, obligat aerobe Bakterien mit einem polaren Flagellum und einer optimalen Wachstumstemperatur von 25-30°C (212). Das namensgebende Charakteristikum (griech. *xanthos*, gelb; *monas*, einzeln) ist die gelbe Färbung der Bakterien, welche durch das membrangebundene Pigment Xanthomonadin bedingt wird und Toleranz gegenüber UV-Strahlung vermittelt (171). Ein weiteres Charakteristikum ist das extrazelluläre Polysaccharid Xanthan, welches adhäsive Eigenschaften besitzt und u.a. als Verdickungsmittel in der Kosmetik- und Lebensmittelindustrie Verwendung findet (8).

Pflanzenpathogene *Xanthomonas* spp. sind hemibiotrophe Pathogene, die lebendes Gewebe kolonisieren und mehr als 120 monokotyledone und 260 dikotyledone Pflanzen infizieren (115). Anhand ihres Wirtsspektrums werden *Xanthomonas* Arten in Pathovare unterteilt. Aufgrund von Ernteverlusten von bis zu 100% gehören die Erreger der Weißblättrigkeit und bakteriellen Streifenkrankheit von Reis, *X. oryzae* pv. *oryzae* (*Xoo*) bzw. *X. oryzae* pv. *oryzicola* (*Xoc*), sowie der Erreger des Bakterienbrandes von Maniok, *X. axonopodis* pv. *manihotis* (*Xam*), zu den wirtschaftlich bedeutsamsten Pflanzenschädlingen (130). Nicht minder relevant sind *X. albilineans* (*Xal*), der Erreger der Blattstreifigkeit von Zuckerrohr, *X. axonopodis* pv. *citri* (*Xac*), der Verursacher des Zitruskrebs' in verschiedenen Zitruspflanzen und *X. campestris* pv. *campestris* (*Xcc*), der Erreger der Adernschwärze von Brassicaceen (130). Zu den etablierten Modellsystemen zur Untersuchung der Interaktion pflanzenpathogener Bakterien mit Wirtspflanzen gehört neben *Xac*, *Xcc* und *Xoo* das in dieser Arbeit untersuchte Pathogen *X. campestris* pv. *vesicatoria* (*Xcv*). *Xcv* wird auch als *X. axonopodis* pv. *vesicatoria* und *X. euvesicatoria* bezeichnet (94,229) und verursacht die bakteriellen Fleckenkrankheit ('bacterial spot disease') auf Paprika (*Capsicum* spp.) und Tomate (*Solanum* spp.)(46,83).

Xanthomonas Bakterien werden durch Regen und Wind im Pflanzenbestand verbreitet und gelangen über natürliche Öffnungen, wie Stomata und Hydathoden, oder Verwundungen in den pflanzlichen Interzellularraum (212). Dort vermehren sich die Bakterien entweder lokal begrenzt, z.B. *Xcv*, *Xac* und *Xoc*, oder verbreiten sich systemisch im Xylem, wie im Falle von *Xcc* und *Xoo* (24). Virulenzfaktoren tragen zur Effizienz und Schwere der Infektion bei, d.h. sind nicht essentiell, wohingegen Pathogenitätsfaktoren für die Vermehrung *in planta* unentbehrlich sind. Ein gut untersuchter Virulenzfaktor von *Xanthomonas* spp. ist Xanthan, welches Bakterienzellen vor Umwelteinflüssen schützt, die Ausbildung von Biofilmen auf der Blattoberfläche und in der Pflanze fördert und zudem zur Ausprägung von Krankheitssymptomen beiträgt (24). Eine Rolle in der Anheftung an Blattoberflächen wurde u.a. für das Adhäsin XadA1 von *Xoo* sowie für Typ IV Pili (Tfp) von *X. campestris* pv. *hyacinthi* beschrieben (172,225). Tfp bestehen aus einem membranverankerten Multi-Proteinkomplex und einem retraktilen Pilus, welcher eine kriechende bzw. gleitende Fortbewegung der Bakterienzelle vermittelt ('twitching/ gliding motility')(91). Studien an *Xoc* und *Xoo* lassen vermuten, dass Tfp zur lokalen bzw. systemischen Ausbreitung der Bakterien im Wirtsgewebe

beitragen (42,240). Virulenzfunktionen wurden außerdem für extrazelluläre Enzyme aus *Xcv*, *Xcc* und *Xoo* beschrieben, z.B. Zellulasen, Endoglucanasen und Xylanasen (24,213). Es wird vermutet, dass solche und andere bakterielle Enzyme am Abbau der pflanzlichen Zellwand beteiligt sind (24).

Untersuchungen an *Xcc* ergaben, dass die Synthese von Virulenzfaktoren, wie Xanthan und extrazellulären Enzymen, mit zunehmender Populationsdichte ansteigt, wohingegen eine geringe Zelldichte die Ausbildung von Biofilmen begünstigt (47,214). Die Regulation der bakteriellen Genexpression in Abhängigkeit von der Populationsdichte wird als ,quorum sensing' (QS) bezeichnet und beruht in *Xcc* auf einer diffussionsfähigen α,ß-ungesättigten Fettsäure (,diffusible signal factor'; DSF)(241). Das *rpf* (,regulation of pathogenicity factors')-Genclusters kommt in allen *Xanthomonas* spp. vor und kontrolliert in *Xcc* die Synthese (RpfF, RpfB) und Perzeption (RpfC, RpfG) von DSF (48). Die extrazelluläre Akkumulation von DSF induziert vermutlich die RpfG-vermittelte Hydrolyse des intrazellulären Botenmoleküls zyklisches di-GMP und fördert dadurch die Synthese extrazellulärer Enzyme (48,78).

Die Pathogenität von *Xanthomonas* spp. sowie der meisten Gram-negativen pflanzen- und tierpathogenen Bakterien beruht auf dem Typ III Sekretionssystem (T3S System), welches Effektorproteine über beide bakterielle Membranen und die pflanzliche Zellwand bzw. Zellmembran in die Wirtszelle transloziert (24,75). Eine Ausnahme innerhalb der Gattung *Xanthomonas* ist *Xal*, für dessen Pathogenität das sekretierte Toxin Albicidin essentiell ist (11,165). Das T3S System pflanzenpathogener *Xanthomonas* spp. wurde erstmals in *Xcv* identifiziert und wird im folgenden Kapitel näher betrachtet (16).

1.4.2. *Xanthomonas campestris* pv. *vesicatoria*

Xcv wird vor allem durch Spritzwasser verbreitet, dringt über Stomata und Wunden in den pflanzlichen Interzellularraum ein und vermehrt sich in anfälligen (suszeptiblen) Pflanzen lokal begrenzt zu hohen Zelldichten (168,206). Die durch *Xcv* verursachte bakterielle Fleckenkrankheit tritt insbesondere in subtropischen und tropischen Regionen auf und ist durch wässrige Läsionen an Blättern und Früchten gekennzeichnet, welche später nekrotisch werden und hohe Ernteverluste verursachen (Abb. 3A und 3B)(95).

Als einer der ersten Vertreter der Gattung wurde im Jahr 2005 die Genomsequenz des *Xcv* Stamms 85-10 veröffentlicht (215). Das Genom besteht aus einem zirkulären Chromosom (~5,18 Mbp) und vier Plasmiden (pXCV2, pXCV19, pXCV38 und pXCV183; 2-183 Kbp) und weist einen für die Gattung charakteristischen G+C Gehalt von 64,75% für das Chromosom und 56 bis 73% für die Plasmide auf. Insgesamt wurden 4.726 ORFs annotiert, welche 87,13% des Genoms ausmachen (215). Biologische Funktionen wurden etwa 65% der ORFs zugewiesen. Die übrigen ORFs kodieren hypothetische Proteine mit unbekannten Funktionen. Das *Xcv* Genom weist zwei rRNA Operons auf, welche jeweils die 16S, 23S und 5S rRNA enthalten, sowie 56 Gene für tRNAs, von denen 54 im Chromosom

lokalisiert sind (215). Außer rRNAs und tRNAs waren zu Beginn dieser Arbeit keine ncRNAs in *Xcv* und anderen *Xanthomonas* spp. bekannt.

Das *Xcv* Genom weist große Ähnlichkeit zu den Genomen von *Xac*, *Xcc* und *Xoo* auf, wobei 66,8% der vorhergesagten Proteine dieser Stämme konserviert sind (40,112,170,215). Die Genome von *Xanthomonas* spp. unterscheiden sich vor allem hinsichtlich des Plasmidgehalts sowie in DNA Regionen, welche einen niedrigen G+C Gehalt aufweisen und meist von IS Elementen flankiert sind (215). Solche Sequenzregionen wurden vermutlich durch horizontalen Gentransfer erworben und kodieren häufig Typ III Effektorproteine (215). Zu den hochkonservierten Bereichen von *Xanthomonas* Genomen gehören u.a. das *rpf*-Gencluster, welches vermutlich an der Synthese und Produktion von DSF beteiligt ist, das *gum*-Gencluster, welches die Xanthanproduktion vermittelt, sowie das *hrp* (‚hypersensitive response and pathogenicity‘)-Gencluster (215).

Das 23-Kbp *hrp*-Gencluster in *Xcv* kodiert das T3S System und ist essentiell für die bakterielle Vermehrung und die Ausbildung von Krankheitssymptomen in suszeptiblen Pflanzen sowie für die Induktion der hypersensitiven Reaktion (HR) in resistenten Pflanzen (16). Die HR ist eine schnelle und lokal begrenzte Zelltodreaktion, welche die weitere Vermehrung des Pathogens verhindert (68,96). Das *Xcv hrp*-Gencluster umfasst 25 Gene, die in acht Transkriptionseinheiten organisiert sind (16,23,27,246). Die Expression des T3S Systems wird in der Pflanze oder im synthetischen XVM2 Medium durch die Schlüsselregulatoren HrpG und HrpX transkriptionell induziert (Abb. 3C)(189,252,253,255). HrpG gehört zur OmpR-Familie der ‚response regulators‘ und wird vermutlich unter *hrp*-Gen-induzierenden Bedingungen posttranslationell aktiviert (254,255). Die verantwortlichen pflanzlichen Signale und *Xcv* Signalproteine sind bislang unbekannt (254). Das aktive HrpG Protein induziert die Transkription von *hrpX*, welches einen Transkriptionsaktivator der AraC-Familie kodiert (252,254,255). Das HrpG-/ HrpX-Regulon, im Folgenden als *hrp*-Regulon bezeichnet, umfasst u.a. das *hrp*-Gencluster, Effektorgene und vorhergesagte Virulenzgene (Abb. 3C)(150,191,215,216,252). Die Transkription der meisten dieser Gene wird durch Bindung von HrpX an ein konserviertes Promotormotiv (PIP Box; ‚plant inducible promoter‘; Konsensus TTCG-N_{16}-TTCG) induziert (105). Grundlage für die funktionelle Charakterisierung des T3S Systems war die Identifizierung einer konstitutiv aktiven HrpG Punktmutante (HrpG*), welche die konstitutive Expression des *hrp*-Regulons unter nicht-induzierenden Bedingungen vermittelt, z.B. in NYG (‚nutrient-yeast-glycerol‘) Komplexmedium (254). Allerdings erfordert die *in vitro* Sekretion von Effektorproteinen spezifische Bedingungen (Minimalmedium A, pH 5.2)(180).

Der Basalapparat des T3S Systems durchspannt beide bakterielle Membranen und wird vermutlich von Hrc (‚*hrp*-conserved‘)-Proteinen gebildet, welche in pflanzen- und tierpathogenen Bakterien konserviert sind (22). Der extrazelluläre Hrp-Pilus dient als Transportkanal für bakterielle Proteine, durchdringt die pflanzliche Zellwand und ist mit bakteriellen Translokonproteinen verbunden, die eine Pore in der pflanzlichen Membran bilden (Abb. 3C)(25,26). Das Pilusprotein HrpE und das potentielle Translokonprotein HrpF gehören zu den nicht-konservierten Hrp-Proteinen und werden über das T3S

System sekretiert (23,26,247). Darüber hinaus tragen sogenannte Hpa (,*hrp*-associated')-Proteine zur Typ III Sekretion bei (25). *hpa*-Gene fördern die Virulenz von *Xcv*, wohingegen *hrc*- und *hrp*-Gene für die Pathogenität essentiell sind.

Die Hauptsubstrate des T3S Systems sind Effektorproteine, welche in *Xanthomonas* spp. als Xop (,*Xanthomonas* outer protein')- bzw. Avirulenz (Avr)-Proteine bezeichnet werden (24). In suszeptiblen Pflanzen hemmen Effektoren die pflanzliche Basalabwehr und ermöglichen dadurch das bakterielle Wachstum sowie die Ausbildung von Krankheitssymptomen (Abb. 3C)(24). Die pflanzliche Basalabwehr wird durch die Erkennung konservierter Pathogen-assoziierte Moleküle, wie Flagellin oder Elongationsfaktor Tu, induziert und umfasst lokale Zellwandverdickungen sowie die Produktion reaktiver Sauerstoffspezies und antimikrobieller Substanzen (109,152). Bislang wurden 26 Effektoren in *Xcv* Stamm 85-10 identifiziert (190). Beispielsweise spaltet XopD SUMO (,small ubiquitin-relatet modifier')-Modifikationen von pflanzlichen Zielproteinen ab und trägt dadurch zur Modulation der pflanzlichen Genexpression und zum bakteriellen Wachstum in Tomate bei (29,85,103). XopS und XopB unterdrücken die Expression pflanzlicher Abwehrgene und hemmen den Vesikeltransport (191). Zudem wurde nachgewiesen, dass XopS und XopB die Vermehrung von *Xcv* in suszeptiblen Paprikapflanzen des Kultivars ECW (,Early Californian Wonder') fördern (191). Resistente Pflanzen sind in der Lage, Avr-Proteine durch spezifische Resistenzgene bzw. Resistenzproteine zu erkennen (24). Beispielsweise induziert die Erkennung der *Xcv* 85-10 Effektoren AvrBs1 und AvrBs2 in resistenten Paprikapflanzen des Kultivars ECW-10R bzw. ECW-20R eine HR (Abb. 3C), welche die Vermehrung von *Xcv* verhindert (135,178,206).

Neben dem T3S System kodiert das Genom von *Xcv* 85-10 Komponenten für alle weiteren Arten von Sekretionssystemen, die bisher in Gram-negativen Bakterien identifiziert wurden. Diese umfassen das Sec- und TAT-System sowie Sekretionssysteme des Typs I bis VI (24,215). Mit Ausnahme des Typ III und Typ II Sekretionssystems ist die Rolle dieser Sekretionssysteme in der Virulenz von *Xcv* unbekannt. *Xcv* kodiert zwei Typ II Sekretionssysteme, welche als Xps und Xcs Systeme bezeichnet werden (215). Kürzlich wurde gezeigt, dass das Xps-, jedoch nicht das Xcs-System, durch Sekretion der Xylanase XynC zur Virulenz und dem *in planta* Wachstum von *Xcv* beiträgt (213). Des Weiteren wurde für *Xcv* 85-10 nachgewiesen, dass die Aconitase AcnB zur Virulenz, dem *in planta* Wachstum und der Verwertung von Citrat als Kohlenstoffquelle beiträgt und Toleranz gegenüber reaktiven Sauerstoffspezies vermittelt (104).
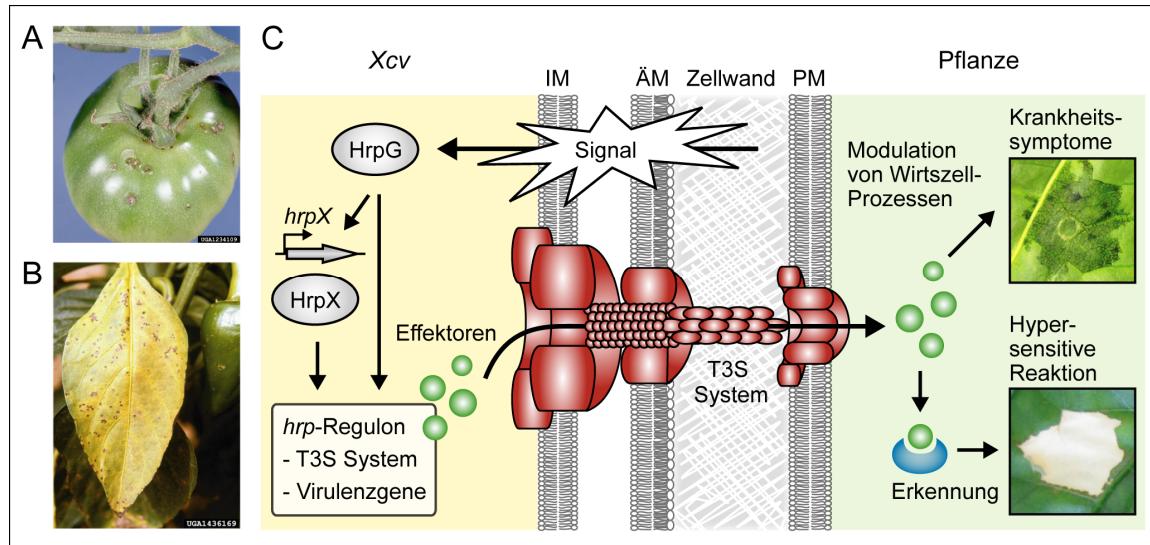
Abbildung 3. Die Interaktion von *Xcv* mit Wirtspflanzen.

Xcv verursacht die bakterielle Fleckenkrankheit auf Früchten und Blättern von (A) Tomate und (B) Paprika. (C) Modell der *Xcv*-Pflanze Interaktion. Im Apoplasten perzipiert *Xcv* unbekannte pflanzliche Signale, welche die Aktivierung von HrpG und die Transkription von *hrpX* induzieren. HrpG und HrpX kontrollieren die Expression eines genomweiten Regulons, welches Gene für Komponenten und Substrate des T3S Systems sowie weitere mögliche Virulenzgene umfasst. Das T3S System vermittelt die Sekretion von Translokonproteinen bzw. die Translokation von Effektoren in die Pflanzenzelle. In suszeptiblen Pflanzen ermöglichen Effektoren die bakterielle Vermehrung. Die resultierenden Krankheitssymptome (wässrige Läsionen) sind als Laborphänotyp auf einem Paprikablatt gezeigt (oben, Kultivar ECW). Die Erkennung von Effektorproteinen durch pflanzliche Resistenzgene bzw. Resistenzproteine induziert die hypersensitive Reaktion, welche als Laborphänotyp auf einem Paprikablatt dargestellt ist (unten, Kultivar ECW-10R). (IM, ÄM: innere und äußere bakterielle Membran; PM: Plasmamembran der Pflanzenzelle. Bildquellen: (A) und (B), Clemson University, USDA Cooperative Extension Slide Series, www.forestryimages.org).

1.5. Zielstellung

Zu Beginn dieser Arbeit waren sRNAs in *Xanthomonas* spp. unbekannt. Ziel dieser Arbeit war die Identifizierung von sRNAs im *Xcv* Stamm 85-10 sowie die funktionelle Charakterisierung ausgewählter Kandidaten in Hinblick auf mögliche Virulenzfunktionen. Vor Beginn dieser Arbeit wurden sRNA Kandidaten durch bioinformatische Analyse der *Xcv* Genomsequenz (215) vorhergesagt, jedoch nicht experimentell validiert (S. Findeiß, F. Thieme, P.F. Stadler und U. Bonas, unveröffentlicht). Die Ergebnisse dieser Analysen sind nicht in die vorliegende Arbeit eingegangen.

Zur Identifizierung von sRNA Kandidaten sollte eine 454-Pyrosequenzierung des *Xcv* Transkriptoms durchgeführt (Kooperation mit C.M. Sharma und J. Vogel) und durch manuelle und bioinformatische Sichtung der Sequenzierdaten analysiert werden (Koop. mit S. Findeiß und P.F. Stadler). sRNA Kandidaten sollten mittels Northern Blot Analysen bestätigt und auf eine HrpG- bzw. HrpX-abhängige Expression getestet werden. Mögliche Virulenzfunktionen von ausgewählten *Xcv* sRNAs sollten durch genetische Analysen wie Deletionsmutagenese sowie durch Wachstums- und Infektionsstudien untersucht werden. Im Falle einer veränderten Virulenz von sRNA Deletionsmutanten sollten die zugrunde liegenden Mechanismen näher charakterisiert werden.

2. Ergebnisse

2.1. Analyse des *Xcv* Transkriptoms

2.1.1. Publikation 1

Genome-wide transcriptome analysis of the plant pathogen *Xanthomonas* identifies sRNAs with putative virulence functions

Cornelius Schmidtke[1,*], Sven Findeiß[2,3], Cynthia M. Sharma[4], Juliane Kuhfuß[1], Steve Hoffmann[3,5], Jörg Vogel[4], Peter F. Stadler[2,3,5,6,7,8,9] and Ulla Bonas[1,*]

[1]Department of Genetics, Martin-Luther-Universität Halle-Wittenberg, Institute for Biology, D-06099 Halle, Germany, [2]Institute for Theoretical Chemistry, University of Vienna, A-1090 Vienna, Austria, [3]Department of Computer Science and Interdisciplinary Centre for Bioinformatics, University of Leipzig, D-04107 Leipzig, [4]Institute for Molecular Infection Biology, University of Würzburg, D-97080 Würzburg, [5]LIFE – Leipzig Research Center for Civilization Diseases, University of Leipzig, D-04107 Leipzig, [6]Fraunhofer Institute for Cell Therapy and Immunology, RNomics Group, [7]Max Planck Institute for the Mathematics in Science, D-04103 Leipzig, Germany, [8]Center for non-coding RNA in Technology and Health, University of Copenhagen, DK-1870 Frederiksberg, Denmark and [9]The Santa Fe Institute, Santa Fe, 87501 New Mexico, USA

ABSTRACT

The Gram-negative plant-pathogenic bacterium *Xanthomonas campestris* pv. *vesicatoria* (*Xcv*) is an important model to elucidate the mechanisms involved in the interaction with the host. To gain insight into the transcriptome of the *Xcv* strain 85–10, we took a differential RNA sequencing (dRNA-seq) approach. Using a novel method to automatically generate comprehensive transcription start site (TSS) maps we report 1421 putative TSSs in the *Xcv* genome. Genes in *Xcv* exhibit a poorly conserved −10 promoter element and no consensus Shine-Dalgarno sequence. Moreover, 14% of all mRNAs are leaderless and 13% of them have unusually long 5′-UTRs. Northern blot analyses confirmed 16 intergenic small RNAs and seven *cis*-encoded antisense RNAs in *Xcv*. Expression of eight intergenic transcripts was controlled by HrpG and HrpX, key regulators of the *Xcv* type III secretion system. More detailed characterization identified sX12 as a small RNA that controls virulence of *Xcv* by affecting the interaction of the pathogen and its host plants. The transcriptional landscape of *Xcv* is unexpectedly complex, featuring abundant antisense transcripts, alternative TSSs and clade-specific small RNAs.

INTRODUCTION

At a staggering pace new high-throughput sequencing technologies have helped to unveil the transcriptional complexity of many organisms in all kingdoms of life (1–3). The recently developed differential RNA sequencing approach (dRNA-seq) has yet added a new perspective. dRNA-seq, based on a selective enrichment of native 5′-ends, has been shown to accurately and cost-effectively identify transcription start sites (TSSs) and RNA processing sites for whole genomes (4). In addition to the obvious advantages for the analysis of 5′-UTR or promoter elements, dRNA-seq allows distinguishing independently transcribed short non-coding and coding RNAs from post-transcriptional processes such as maturation (4). However, a fully-automated method to annotate and statistically evaluate TSSs in large dRNA-seq data sets has been missing so far. Here, we sketch a procedure to automatically identify TSSs.

Transcriptome analyses in plant pathogenic bacteria so far mainly focused on coding regions and the regulon controlling type III secretion [e.g. (5,6)]. A

recent deep sequencing analysis of *Pseudomonas syringae* identified many small RNA (sRNA) candidates, most of which, however, await validation by independent methods (7).

The Gram-negative plant pathogenic γ-proteobacterium *Xanthomonas campestris* pv. *vesicatoria* (*Xcv*) is the causal agent of bacterial spot disease on pepper and tomato and is of great economic importance in regions with a warm and humid climate (8). *Xcv* serves as a model system to elucidate the molecular communication between plant pathogens and their hosts and to characterize bacterial virulence strategies. Genome analysis predicted 4726 open reading frames (ORFs) in the *Xcv* strain 85–10 (9), yet the overall gene structure and non-coding RNA output of this model pathogen are still poorly understood.

Essential for pathogenicity of *Xcv* on susceptible host plants is the type III secretion (T3S) system, encoded by the *hrp* [hypersensitive response (HR) and pathogenicity] gene cluster (10). In *Xcv*, as in most Gram-negative bacterial pathogens, the T3S nanomachine translocates a suite of effector proteins into the plant cell where they manipulate host cellular processes to the benefit of the pathogen, e.g. by suppression of basal plant defense responses (9,11–13). *hrp* mutants do not grow in plant tissue, and they no longer cause disease in susceptible plants and the HR in resistant plants (10). The HR is a local, rapid programmed cell death at the site of infection, which coincides with arrest of bacterial multiplication in the plant (14,15).

The T3S system is transcriptionally induced in certain minimal media and in the plant (16,17). Key regulatory proteins are the OmpR-type response regulator HrpG, which is activated by unknown plant signals and controls the expression of a genome-wide regulon including *hrp*, type III effector and putative virulence genes (16–19). HrpG-mediated activation of gene expression depends in most cases on the AraC-type transcriptional activator HrpX (18), which binds to a conserved motif (plant-inducible promoter; PIP box) in the promoters of target genes (20). The identification of a point mutation in HrpG (termed HrpG*), which renders the protein constitutively active, was key for the analysis of T3S and the identification of putative virulence factors that are cotranscribed with the T3S system (19,21). An open question was whether virulence gene expression in *Xcv* is post-transcriptionally regulated, for instance by sRNAs. Here, we provide for the first time an insight into the transcriptional landscape of a plant pathogenic bacterium and the involvement of sRNAs in its virulence.

MATERIALS AND METHODS

RNA isolation for 454 pyrosequencing, RACE analysis and northern blot

RNA was isolated from NYG-grown *Xcv* strains 85–10 and 85* (exponential growth phase) by phenol extraction and treated with DNase I (Roche). For RACE and northern blot analyses, RNA was isolated from NYG-grown *Xcv* strains in exponential and stationary growth phases, as described (22). RACE analyses were carried out as described (23) with modifications [for detailed information see Supporting Information (SI)]. Northern blots were performed as described (24) using 10 μg RNA, 5–10 pmol [γ-^{32}P]-ATP end-labeled oligodeoxynucleotides (Supplementary Table S1). Hybridization signals were visualized with a phosphoimager (FLA-3000 Series, Fuji). Northern blot hybridizations were performed at least twice with independently isolated RNA.

Construction of cDNA libraries for dRNA-seq and 454 pyrosequencing

Prior to RNA treatment and cDNA synthesis, equal amounts of RNA from the two *Xcv* strains 85–10 and 85* were mixed. dRNA-seq libraries were prepared according to Sharma *et al.* (2010) and sequenced with a Roche 454 sequencer using FLX and Titanium chemistry (see SI).

Annotation of transcription start sites

We aimed at the automated identification of TSSs based on the discrimination between narrow clusters of dRNA-seq reads that might represent a TSS and the distribution of individual read starts. The density of read starts varies across the genome and can be modeled locally by a Poisson distribution with a parameter λ. We used fixed-length intervals of size l to determine $\lambda_r = s_r/l$ from the number of read starts s_r in the region r. The parameter λ_{ave} models the average genome wide arrival rate of read starts. λ is defined as λ_r/λ_{ave}. The corresponding Poisson distribution $F(k,\lambda)$ describes the probability that at most k read starts are observed at a given genomic position. We used library 1 to determine λ_m for the background distribution of read starts. Similarly, library 2 was used to obtain λ_p to model the distribution biased towards the TSS.

A TSS is defined as the genomic position at which the observed number of read starts in library 2 significantly exceeds the background distribution of read starts in library 1. The significance of a putative TSS was determined as follows: for each genomic position, the difference of the number of read starts P in library 2 and M in library 1, $D = P{-}M$, was calculated. The difference of two Poisson distributed variables, D, follows a Skellam distribution (25) whose cumulative distribution function is given by

$$F(D,\lambda_p,\lambda_m) = \sum_{d=-\infty}^{D} e^{-(\lambda_p+\lambda_m)} \left(\frac{\lambda_p}{\lambda_m}\right)^{\frac{d}{2}} J_{|d|}(2\sqrt{\lambda_p\lambda_m}); \; d \in Z$$

where $J_{|d|}$ is the modified Bessel function of the first kind and integer order $|d|$. Furthermore, $1 - F(D,\lambda_p,\lambda_m)$ represents the probability that a difference of at least D read starts is observed given the normalized rates of read starts λ_p and λ_m. To reduce the influence of window sizes and local variation of transcriptional activity a sliding window of size x was shifted by y nucleotides along the genome

and each site was tested $t = x/y$ times for being a TSS. The *p*-value was obtained using the geometric mean

$$p = \sqrt[t]{\prod_{i=1}^{t} p_i}$$

where p_i denotes the *P*-value obtained in the *i-th* test. Note that only sites with a minimum expression of three read starts within a distance of ≤ 5 nt were tested. Furthermore, we excluded sites in the vicinity of perfectly aligned hit blocks, i.e. stacks of hits that all share a common 5′- and 3′-end. To determine λ_r, we selected a region size of 500 nt. For the sliding window approach an offset of 50 nt was used. All potential TSSs significant to the $p = 0.05$ level are listed in Supplementary Table S2. In order to achieve a high positive predictive value for data sets of similar size, these parameters have been fixed globally in our study and may have to be adjusted for the application of the method to other data sets.

Evaluation of the automated TSS annotation method

To evaluate the predictive power of the automated TSS annotation method we used *Helicobacter pylori* and its manually curated TSS map (4) as reference. A data set of comparable size to the *Xcv* data set was generated. Reads overlapping with annotated tRNA or rRNA genes were excluded. From the *H. pylori* data set 40 385 mapped reads of the treated library and 49 845 reads of the untreated library were randomly selected and contained 392 manually annotated TSSs which were used as reference class. TSSs were predicted using the same parameter settings (500 nt window size, 50 nt offset; 0.05 *p*-value cutoff) as for the *Xcv* data set. 566 genomic positions met the criteria for being TSS candidates, i.e. the clustering of at least three read starts. These positions represent putative TSSs and were statistically evaluated with the automatic TSS annotation approach, according to (26). The results are summarized in an extended confusion matrix (Supplementary Table S9).

Estimation of expression level

To estimate the expression level of CDSs in *Xcv* likely to exhibit a proximal promoter, we selected 1276 annotated CDSs in a head-to-head arrangement. The set comprised 549 CDSs with and 727 without annotated TSS. Due to the limited sequencing depth of our data set we combined reads of both libraries and evaluated the coverage of the first 100 nt of CDSs (Supplementary Figure S2).

Detailed information about additional methods is provided in SI.

Further supporting information and the raw sequencing data are available at the official institutional website of the University of Leipzig (http://www.bioinf.uni-leipzig.de/publications/supplements/10-035).

RESULTS

Mapping of sequencing reads

To analyze the primary transcriptome of *Xcv*, total RNA of strain 85–10 and its derivative 85* were mixed (SI and Supplementary Table S1). *Xcv* strain 85* carries a chromosomal point mutation in *hrpG* (*hrpG**) leading to expression of the Hrp-regulon. cDNAs were synthesized from total RNA (untreated library; hereafter library 1) and RNA enriched for primary transcripts (treated library; hereafter library 2), respectively (4). dRNA-seq analysis resulted in 160 349 reads for library 1 and 149 596 reads for library 2. A total of 84% of the reads were mapped to the *Xcv* genome using the program segemehl (27). As previously described, *Xcv* contains two identical copies of the 5S, 23S and 16S rRNA clusters, respectively, and 56 tRNA loci (9). A total of 63% of the reads of library 1 and 68% of library 2 reads mapped to these genes although the processed rRNAs and tRNAs were expected to be depleted in library 2. Closer examination revealed that the majority of tRNA-read starts in library 2 correspond to the presumed RNase P processing sites rather than TSSs (Supplementary Figure S1). To verify our observations we analyzed all reads overlapping tRNAs in the *Helicobacter pylori* dRNA-seq data set (4), which supports our findings (Supplementary Figure S1). The abundance of library 2 tRNA reads mapping to putative RNase P processing sites might be due to stable secondary structures formed after RNase P cleavage thus protecting mature tRNAs from exonuclease degradation. We, therefore, discarded the reads mapping to rRNA and tRNA loci and analyzed the remaining 49 845 and 40 385 reads in more detail. While reads of library 1 cover entire genes, the read starts of library 2 are shifted towards the 5′-end of primary transcripts, which permits precise mapping of the TSS of a given gene (Figure 1A, e.g. *XCV0520*), as described (4).

A statistical model to annotate TSSs

Most of the TSS maps published to date are derived from tedious manual inspection of sequencing data (4,24,28) or using *ad hoc* heuristics complemented by manual inspection (29–31). Here, we aimed at the automated identification of TSSs based on well-defined criteria, i.e. to discriminate between potential TSSs and the background distribution of read starts. This background, however, is not uniform across the genome but varies depending on gene expression levels. We therefore modeled read starts by Poisson distributions depending on the expression level in a well-defined genomic neighbourhood. Comparing the two libraries, a TSS is defined as a position where the observed difference of read starts in both libraries significantly exceeds the expected differences of read starts modeled by a Skellam distribution from which *p*-values are readily derived (see 'Materials and Methods' section).
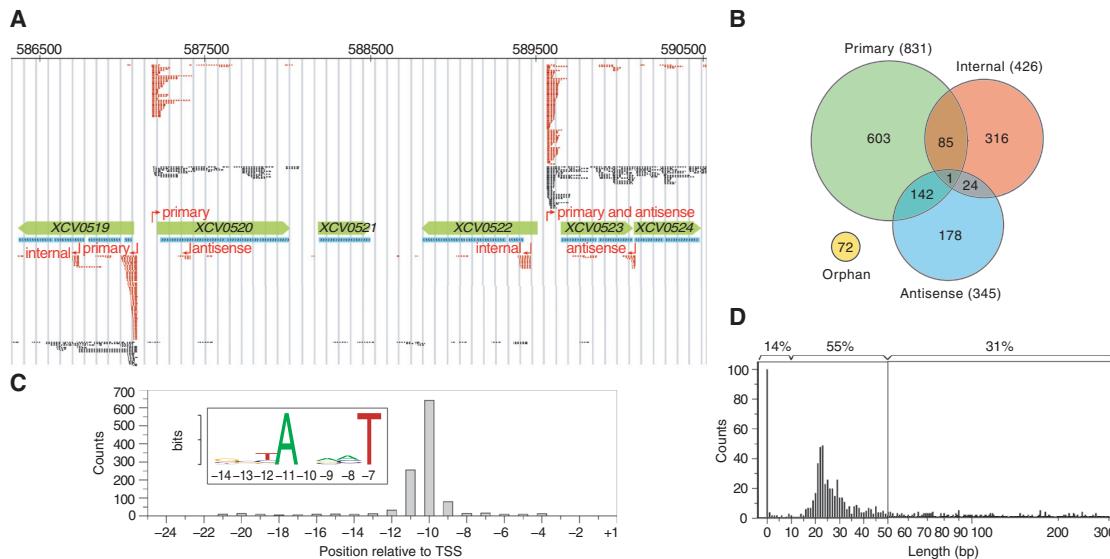
Figure 1. Identification of TSSs, promoter elements and analysis of 5′-UTRs. (**A**) Distribution of dRNA-seq reads in the chromosomal locus of *Xcv* 85–10 spanning genes *XCV0519* to *XCV0524*. Annotated CDSs and RNAcode high-scoring segments are highlighted in green and blue, respectively. Sequencing reads of library 1 (black) and library 2 (red) are shown on top for the (+)-strand and below for the (−)-strand. Predicted TSSs and corresponding classes are indicated in red. (**B**) Venn diagram illustrating the TSS classes. TSSs found maximal 300-bp upstream of coding sequences are classified as primary. Internal TSSs are found within and antisense TSSs on the opposite strand of genes (± 100 bp). Orphan TSSs do not belong to other classes. (**C**) Sequence analysis identified a T/A-rich promoter element for 1205 of 1421 putative TSSs. The histogram depicts the position of the conserved sequence pattern relative to the annotated TSSs at position +1. (**D**) 5′-UTR length distribution. The *x*-axis is split into linear (0–50) and logarithmic (51–300) scales. The top of the histogram gives the percentage of leaderless (≤10 bp), short (≤50 bp) and longer UTRs (>50 bp).

Annotation of TSSs

In total, 1372 chromosomal and 49 TSSs on the large plasmid pXCV183 of *Xcv* (Figure 1B and Supplementary Table S2) were identified. The data confirm TSSs determined previously for selected pathogenicity genes, e.g. *hrcU* and *hrpB1* (20,32). Nevertheless, the majority of TSSs annotated in our study should be considered as putative. TSSs were classified into four categories, i.e. (i) primary TSSs located up to 300 bp 5′ of an annotated translation start, (ii) internal TSSs within an annotated coding sequence (CDS), (iii) antisense TSSs that map to the opposite strand of CDSs ± 100 bp and (iv) orphan TSSs that do not belong to the other three categories. Most of the annotated TSSs are primary TSSs (831) and probably correspond to the 5′-end of mRNAs. Overall, CDSs that lack an assigned TSS exhibit much lower expression levels than CDSs with an annotated TSS (see 'Materials and Methods' section and Supplementary Figure S2).

As illustrated in Figure 1B, TSSs can belong to more than one category, e.g. the assumed primary TSS of *XCV0523* is also antisense to *XCV0522* (Figure 1A). Interestingly, 10% (86/831) of primary TSSs are also classified as internal. Thus, some neighboring CDSs previously supposed to be cotranscribed as part of a polycistronic mRNA can also be transcribed from alternative promoters. As illustrated for *XCV0522* (Figure 1A), we

identified 71 putative TSSs which are located within the first 50 bp of annotated CDSs suggesting that previously annotated translation starts have to be revisited (Supplementary Table S3). Furthermore, 345 TSSs are located antisense to annotated genes. Interestingly, 41% of these TSSs are also classified as primary TSSs, including 16 TSSs that correspond to overlapping mRNAs in an antisense orientation (Supplementary Table S2). 49 antisense TSSs are positioned in the 3′-region (± 100 bp) of annotated sense genes (Supplementary Table S4). In total, antisense reads map to 22% of all nucleotides that belong to annotated CDSs irrespective of read numbers, the presence of a TSS and the expression of the corresponding CDSs. The majority of these antisense reads lack automatically assigned TSSs and do not accumulate in clusters and thus, might not be originated from defined antisense genes. We also compared the sense- and antisense-read coverage of all annotated CDSs in *Xcv* and did not observe a correlation (data not shown).

Most bacterial δ^{70}-dependent promoters contain conserved sequence elements, i.e. −35 (TTGACA) and −10 (TATAAT) elements present in *Escherichia coli* (33). In *Xcv*, there is a weakly conserved T/A-rich motif in the proximity of −10 regions, however, other conserved promoter elements and a Shine-Dalgarno (SD) motif are missing (Figure 1C). This might be due to the high G + C content (65%) of the *Xcv* genome (9) and is discussed below.
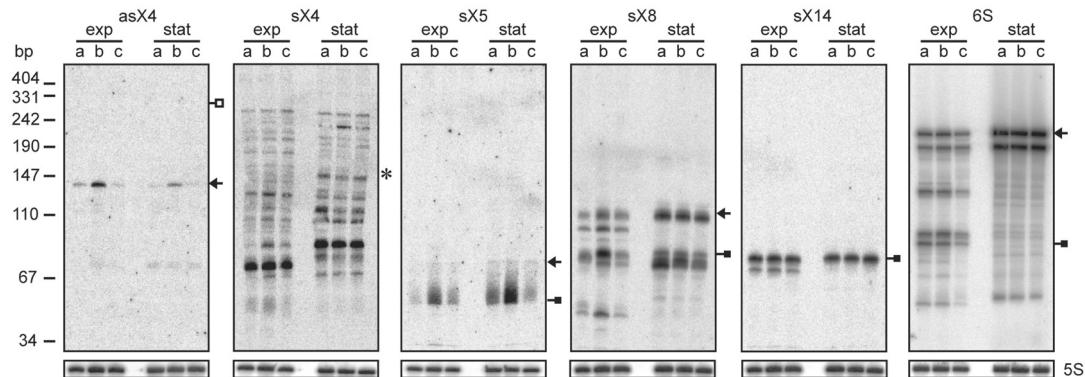
Figure 2. Expression of selected *Xcv* sRNAs and antisense RNAs depends on HrpG and HrpX. Total RNA isolated from exponential (exp) and stationary phase cultures (stat) of (a) *Xcv* strain 85–10, (b) 85–10 expressing *hrpG** from pFG72-1 and (c) 85–10Δ*hrpX* carrying pFG72-1 was analyzed by northern blot. Arrows and filled squares indicate signals corresponding to the expected full-length RNA and processing products obtained by transcriptome sequencing, respectively. The open square indicates the expected size of full-length asX4 determined by RACE analysis. The expected size of sX4 according to the sequencing data is marked by an asterisk. 5S rRNA (lower panel) was probed as loading control.

Analysis of 5′-UTRs revealed unexpected size diversity

The lengths of 5′-UTRs deduced from 831 putative primary TSSs range from 0 to >300 bp, with the majority being between 10 and 50 bp (Figure 1D). Surprisingly, 14% of the mRNAs (118 of 831) are leaderless, i.e. their 5′-UTR consists of <10 bp with respect to the annotated genome sequence of *Xcv* (9). Many of the corresponding genes presumably have housekeeping functions (Supplementary Table S5). In addition, the 5′-UTRs of type III effectors were manually inspected. TSSs of 11 described type III effectors from *Xcv* strain 85–10 (9,13) were mapped in this study (Supplementary Table S5). The promoter regions of nine effector genes contain a PIP box (consensus TTCG-N$_{16}$-TTCG) (20). The assumed lengths of the 5′-UTRs of *avrBs2*, *xopE2*, *xopJ1* and *xopO* are average. Curiously, the *avrRxv* mRNA is leaderless, and six mRNAs (*avrBs1*, *xopAA*, *xopB*, *xopC*, *xopD* and *xopN*) (9,13) contain unusually long 5′-UTRs, ranging from 173 to 678 bp. Consequently, the CDSs of some effector genes might be considerably larger than previously described (9). Overall, 13% of the *Xcv* 5′-UTRs are unusually long (150–300 bp; Supplementary Table S5).

Northern blot analysis confirmed 23 sRNAs in *Xcv*

A computational scan for known RNA elements in *Xcv* identified already annotated tRNAs, rRNAs and the recently described ptaRNA1 (34). In addition, we identified eight putative riboswitches and widely conserved RNAs, i.e. RNase P-, RtT-, SRP-, tmRNA and 6S-RNA (Figure 2 and Supplementary Table S6). Based on our dRNA-seq data, most of these transcripts were strongly expressed and TSSs were annotated for four of the housekeeping RNAs and five of the predicted riboswitches (Supplementary Table S6). The genes located downstream of the riboswitch candidates are either known to be involved in the respective riboswitch-controlled pathways in other bacteria or, as in case of yybP/ykoY candidates, presumably encode membrane proteins (35–37) (Supplementary Table S6).

Prior to the automated TSS prediction, we selected 89 sRNA candidates by manual inspection of the sequencing data with a focus on intergenic regions. We used northern blot analysis to experimentally validate sRNA candidates and analyzed their potential coregulation with the T3S system. To this end, RNA was isolated from exponential and stationary phase cultures of NYG-grown *Xcv* strains 85–10, 85–10 expressing HrpG* and a derivative lacking *hrpX* (85–10Δ*hrpX*p*hrpG**), respectively. Northern hybridizations confirmed 23 new sRNAs, whereas remaining candidates either appeared to correspond to longer transcripts, i.e. UTRs of mRNAs, or were poorly detectable. The latter can be explained by their low abundance in the dRNA-seq data (data not shown).

After completion of bioinformatic analyses, seven verified sRNAs turned out to correspond to *cis*-encoded antisense RNAs, termed asX1-7 (Table 1, Figure 2 and Supplementary Figure S3). We detected dRNA-seq reads mapping to both antisense RNA and mRNA for six of these transcripts and a few reads mapping to the CDS complementary to *asX4*, respectively (data not shown). The remaining 16 sRNAs mapped to intergenic regions and were termed sX1-15 and 6S (Table 1, Figures 2, 3A, 4A and Supplementary Figure S3). Intriguingly, three sRNAs (sX15, asX6, asX7) are encoded on the large plasmid, two of which (asX7 and sX15) are in antisense orientation to each other (Table 1 and Supplementary Figure S3). Most sRNA genes were constitutively expressed under the conditions tested, and appeared to accumulate in stationary growth phase either due to higher transcription rates or increased stability, e.g. sX14 and 6S (Figure 2). Interestingly, expression/accumulation of five intergenic sRNAs and three antisense RNAs was affected by the key regulators of *hrp* gene expression, HrpG and HrpX, suggesting a role of these sRNAs or their targets in the interaction of *Xcv* with the plant. HrpX-dependent induction of sRNA expression was observed for asX4, sX5, sX8 (Figure 2) and sX12 (see below), whereas sX11 appeared to be HrpG/HrpX-dependently repressed

Table 1. Verified sRNAs (sX) and antisense RNAs (asX) in *Xcv*

RNA (Strand)	TSS category[a]	Start-Stop[b]	Library 2[c]	Library 1[c]	Expected length (nt)[d]	Detected length (nt)[e]	HrpG/HrpX dependency[f]	Conservation[g]
sX1 (−)	primary: *XCV0067*	**78978**–78799	8	5	180	190	–	A (5); B (4); C1 (4); C2 (5);
		78978–78816	8	5	163	170		C3 (6); D1–3 (4); E (3)
		78978–78937	8	5	42	50		
sX2 (+)	orphan	**85087**–85196	11	0	110	110	–	A; B; C1–3; D1–3
		85109–85196	0	5	88	85		
sX3 (+)	orphan	**1233578**–1233669	28	5	92	85	–	A
		1233611–1233669	0	14	59	60		
sX4 (+)	orphan	**1235373**–1235528	5	1	156	280–300	stability	A; B; C1–3
sX5 (−)	–	1899107–1899037	2	0	71	70	HrpX–induced	A; B; C1–3; D1–3; G1–2
		1899085–1899037	35	41	49	50		
sX6 (+)	antisense: *XCV1748*	**1971505**–1971845	15	1	341	350	–	A; B; C1–3; D1–3
sX7 (+)	–	1995660–1995754	1	57	95	85	–	A; B; C1–3; D1–3; E
sX8 (+)	orphan	**2740875**–2740992	7	3	118	110	HrpX–induced	A; F1; H; I; J
		2740875–2740956	7	3	82	85		
sX9 (+)	–	2929816–2929890	2	1	75	75	–	A; B
sX10 (+)	–	3850318–3850496	25	17	179	180	–	A (5); B (4); C1 (4); C2 (5);
		3850415–3850496	14	6	82	80		C3 (6); D1–3 (4); E (3)
sX11 (+)	orphan	**4069950**–4070087	35	10	138	130	HrpG/HrpX–repressed	A; B; C1–3; D1–3
		4070000–4070087	0	5	88	100		
sX12 (+)	–	<u>4358796</u>–<u>4358873</u>	4	2	78	67	HrpX–induced	A; B; C1–3; D1–3
sX13 (−)	–	4810196–4810082	2	73	115	105	–	A; B; C1–3; D1–3; E; G1–2
sX14 (−)	–	5040690–5040599	2	16	92	85	–	A; B; C–3; D1–3
sX15 (+)	antisense: *XCVd0106* and primary: *XCVd0107*	#**116378**–116536	80	8	159	150	–	Ap; G1 (10); G2
		#116438–116536	2	4	99	100		
6S (+)	orphan	**4037865**–4038084	2627	442	220	220	–	A; B; C1–3; D1–3; E; F1–4;
		4037865–4037952	2627	442	88	90		G1–2
asX1 (+)	–	447108–447223	2	0	116	110	stability	A; B; C1–3; D1–3;E; F1–4; G1–2
asX2 (−)	orphan	**3290997**–3290913	12	2	85	75	–	A; B
asX3 (+)	orphan	**4498825**–4499001	16	0	177	70–600	–	A; D1–3
asX4 (−)	antisense: *XCV4106*	**4701994**–4701856	23	7	139	140	HrpX–induced	A; B; C1–3; D1–3
		4701994–<u>4701686</u>	23	7	309	–		
asX5 (+)	orphan	**4757260**–4757445	21	2	186	190	stability	A
		4757360–4757445	1	7	86	85		
asX6 (+)	antisense: *XCVd0099*	#**109851**–109943	35	6	93	90	–	Ap; Bp; Kp; Lp
		#**109851**–109875	35	6	25	30		
asX7 (−)	–	#**116528**–116459	4	4	70	60	–	Ap; G1 (10); G2

[a]Classification of the automatically annotated TSS (Figure 1B). [b]The 5′- and 3′-positions of the respective dRNA-seq-read clusters on the *Xcv* chromosome and plasmid pXCV183 (indicated by #). Positions highlighted in bold indicate an automatically annotated TSS (see SI; Supplementary Table S2). Underlined numbers correspond to transcript ends which were verified by 5′- and 3′-RACE, respectively. The underlined 3′-end of asX4 was identified only by RACE. [c]Number of read starts at the respective start position given in column 'Start-Stop'. [d]Transcript length deduced from dRNA-seq. [e]transcript size and [f]HrpG/HrpX dependency of sRNA/antisense RNA accumulation determined by northern blot (Figures 2, 3, 4 and Supplementary Figure S3); 'stability' indicates altered amounts of sRNA processing products in dependency of HrpG and/or HrpX. Constitutive expression is indicated by '-'. [g]sequence conservation among other bacteria (see SI). Strains containing homologous sequences and the respective accession numbers are given below. Numbers in brackets indicate the number of homologous sequences in the respective strains if more than one homolog was identified.

A: *X. campestris* pv. *vesicatoria* 85–10 (NC_007508)
Ap: *X. campestris* pv. *vesicatoria* 85–10 plasmid pXCV183 (NC_007507)
B: *X. axonopodis* pv. *citri* 306 (NC_003919)
Bp: *X. axonopodis* pv. *citri* 306 plasmid pXAC64 (NC_003922)
C1: *X. campestris* pv. *campestris* ATCC 33913 (NC_003902)
C2: *X. campestris* pv. *campestris* 8004 (NC_007086)
C3: *X. campestris* pv. *campestris* B100 (NC_010688)
D1: *X. oryzae* pv. *oryzae* MAFF 311018 (NC_007705)
D2: *X. oryzae* pv. *oryzae* KACC 10331 (NC_006834)
D3: *X. oryzae* pv. *oryzae* PXO99A (NC_010717)
E: *X. albilineans* GPE PC73 (NC_013722)
F1: *Xylella fastidiosa* 9a5c (NC_002488)
F2: *Xylella fastidiosa* Temecula1 (NC_004556)
F3: *Xylella fastidiosa* M12 (NC_010513)
F4: *Xylella fastidiosa* M23 (NC_010577)
G1: *Stenotrophomonas maltophilia* K279a (NC_010943)
G2: *S. maltophilia* R551-3 (NC_011071)
H: *Burkholderia xenovorans* LB400 (NC_007951)
I: *Acidovorax* sp. JS42 (NC_008782)
J: *Bordetella petrii* DSM 12804 (NC_010170)
Kp: *Ralstonia solanacearum* CMR15 plasmid pRSC35 (FP885893)
Lp: *X. citri* plasmid pXcB (AY228335)

(Supplementary Figure S3). In case of sX4 (Figure 2) and the antisense RNAs asX1 and asX5 (Supplementary Figure S3) the sRNA stability appeared to depend on HrpG and HrpX as well as on the growth phase.

Processing of sRNAs

In general, the dRNA-seq data and northern blots suggest that *Xcv* sRNAs do not accumulate as primary transcripts
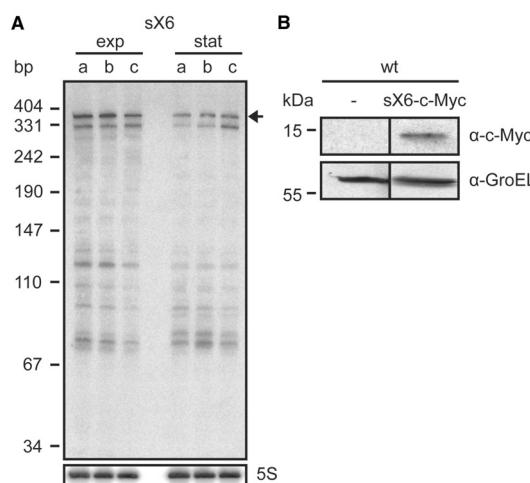
Figure 3. sX6 encodes a small protein. (**A**) Expression analysis of the sX6 transcript. Total RNA isolated from exponential (exp) and stationary phase cultures (stat) of (a) *Xcv* strain 85–10, (b) 85–10 expressing *hrpG** from pFG72-1 and (c) 85–10Δ*hrpX* carrying pFG72-1 was analyzed by northern blot. The expected signal according to sequencing data is indicated by an arrow. 5S rRNA (lower panel) was probed as loading control. (**B**) Expression of the sX6 protein. Derivatives of *Xcv* strain 85–10 (wt) carrying promoterless empty vector pBRM-P (−) and *sX6*-c-Myc expression construct, respectively, were grown to $OD_{600} = 0.7$. Protein extracts were analyzed by immunoblotting using c-Myc epitope-specific and GroEL-specific antibodies.

but undergo growth-phase dependent processing. However, in most cases the apparent sizes of full-length and processed sRNAs in northern blots were in agreement with the dRNA-seq data, e.g. sX8 and 6S RNA (Figure 2 and Table 1). In addition to full-length and processing products, northern blots detected unexpectedly long signals, up to 900 nt, for the antisense RNAs asX1, asX2, asX3, asX6 and asX7 (Supplementary Figure S3). These signals may be caused by alternative termination of transcription. The sequencing data also suggest that sX7, sX13 and sX14 represent processing products of longer transcripts since reads mapping to these loci are predominantly found in library 1, and no TSS was identified in library 2 (Table 1). For selected RNAs the 5′- and 3′-ends were determined by RACE (Table 1). While the 5′-end of the antisense RNA asX4 is identical to the TSS identified by dRNA-seq, the 3′-region is 170 nt longer suggesting the presence of a processing site.

Phylogenetic distribution of sRNAs from *Xcv*

While *sX3* and *asX5* are unique for *Xcv*, homology searches revealed that 10 sRNA genes are exclusively found in sequenced *Xanthomonas* species that encode a *hrp*-T3S system (Table 1). Four of the latter sRNAs, including sX12 described in more detail below, and asX5 were coregulated with the T3S system.

Two intergenic sRNAs, sX1 and sX10 (Table 1; Supplementary Figure S3) are highly similar in sequence and structure. Three additional homologous genes are predicted and expressed in *Xcv* and might therefore be considered as an sRNA family. As three to six copies of members of this gene family are found in other *Xanthomonas* species (Table 1), we propose a functional redundancy of the respective sRNAs.

Interestingly, 10 homologs of the plasmid-encoded and complementary *Xcv* sX15 and *asX7* genes are present in the chromosome of *Stenotrophomonas maltophilia* strain K279a (38) (Table 1). Moreover, *asX6*, which is also
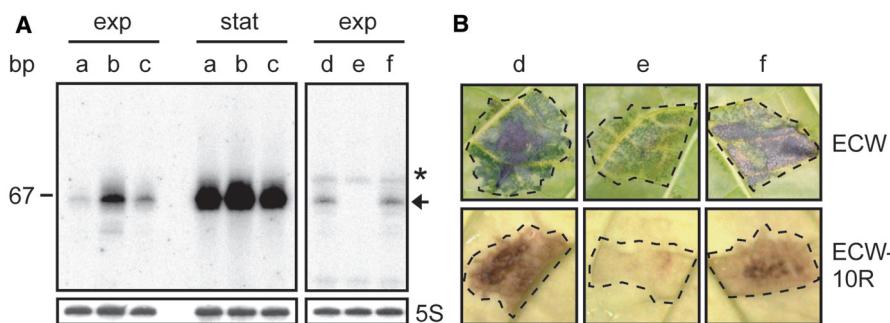
Figure 4. *sX12* is involved in virulence of *Xcv*. (**A**) *sX12* is HrpX-dependently expressed. Total RNA isolated from exponential (exp) and stationary phase cultures (stat) of (a) *Xcv* strain 85–10, (b) *Xcv* expressing *hrpG** from pFG72-1 and (c) a derivative deleted in *hrpX* and carrying pFG72-1 was analyzed by northern blot. The right panel shows a northern blot with RNA from (d) *Xcv* strain 85–10 and (e) an *sX12* deletion mutant carrying empty vector pLAFR6, respectively, and (f) an *sX12* deletion mutant ectopically expressing *sX12* from p*sX12*. The expected RNA size is indicated by an arrow. The asterisk denotes an unspecific signal. 5S rRNA (lower panel) was probed as loading control. (**B**) *sX12* contributes to virulence and the HR. Strains used in (A) (right panel) were inoculated at a density of 1.25×10^8 CFU ml^{-1} into leaves of susceptible ECW and resistant ECW-10R pepper plants. Disease symptoms were photographed at 7 days post-inoculation (dpi). The HR was visualized by ethanol bleaching of the leaves at 2 days post-inoculation. Dashed lines indicate the inoculation site.

located on pXCV183 of *Xcv*, is conserved in plasmids of *X. axonopodis* pv. *citri* strain 306 (39), *Ralstonia solanacearum* strain CMR15 (40) and *X. citri* (41) (Table 1). A rather erratic phylogenetic distribution was observed for *sX8* since homologs are predicted in a small subset of the known genomes of both beta- and gamma-proteobacteria (Table 1). Interestingly, this holds true also for the gene cluster upstream of *sX8* which suggests a common evolutionary origin of this region. This type of phylogenetic pattern has in particular been observed for toxin/anti-toxin systems and suggests frequent horizontal transmissions (42).

sX6 encodes a small protein

Using RNAcode, a program, which was applied for the detection of novel protein coding genes in *E. coli* (43), 24 potential short ORFs were predicted in the *Xcv* genome (see SI and Supplementary Table S7). dRNA-seq reads mapped to 12 of these loci. One example is sX6 (341 nt), which is constitutively expressed (Figure 3A) and has a predicted coding capacity of 80 amino acids including a signal peptide in the N-terminal region. We generated a translational fusion of sX6 with a C-terminal c-Myc epitope tag, under control of the native *sX6* promoter (44), and introduced the expression construct into *Xcv* strain 85–10. As shown in Figure 3B, a fusion protein of the predicted molecular mass ($\sim$12 kDa) was detectable in protein extracts of *Xcv*.

Besides *sX6*, TSSs for two of the predicted ORFs with a coding capacity of 36 and 67 amino acids, respectively, were predicted (Supplementary Table S7). Interestingly, homologs of genes for the three small proteins are exclusively found in xanthomonads encoding a *hrp*-T3S system.

sX12 contributes to virulence

The fact that several sRNAs are expressed under control of the T3S system regulators suggested a possible role in virulence. Here, we focused on sX12 whose size of 78 nt was confirmed by 5'- and 3'-RACE (Table 1). As mentioned above, expression of *sX12* is HrpX-dependently induced and accumulates in stationary growth phase (Figure 4A). To assess the contribution of *sX12* to virulence we generated a deletion mutant derivative of strain 85–10 ($\Delta sX12$). While growth of strain $\Delta sX12$ *in planta* was as wild-type (Supplementary Figure S4), plant reactions were altered. Disease symptoms in leaves of infected susceptible (ECW) and the HR in resistant (ECW-10R) pepper plants were delayed with strain $\Delta sX12$ when compared to the wild-type (Figure 4B). The $\Delta sX12$ mutant phenotype was complemented by ectopic expression of *sX12* under control of its own promoter (Figure 4). We also performed T3S assays to analyze whether the delay in plant reactions by strain $\Delta sX12$ might be due to reduced protein levels of T3S system components, e.g. the conserved apparatus component HrcJ, or the secretion of T3S substrates, i.e. the translocon protein HrpF. However, the detected protein amounts and the secretion of HrpF were comparable for the wild-type and the $\Delta sX12$ mutant (Supplementary Figure S4).

DISCUSSION

The dRNA-seq-based analysis of the *Xcv* transcriptome led to remarkable insights into the transcriptional landscape of this important model plant pathogen and identified an sRNA with a role in virulence. In this study, we have devised a new method to automatically generate maps of TSSs for dRNA-seq data sets alleviating the need for manual inspection and allowing application of dRNA-seq also for larger genomes than *Xcv*. In contrast to earlier dRNA-seq approaches, mostly based on laborious manual inspection of sequencing data (4,24,28), the presented computational approach provides a measure of statistical confidence and ensures that predictions are comparable between different studies as demonstrated by our comparative analysis between manual and automated annotation of the previously published *H. pylori* transcriptome (4). While the sensitivity of 82% demonstrates the method's capability of recovering manually annotated TSS at exactly the same position, a positive predictive value of 72% indicates its reliability (Supplementary Table S9). However, to dynamically adjust parameters such as significance levels the method remains subject to further research. We used only exact matches of the manual and automated TSS map for this analysis. The number of false positives and negatives might therefore be overestimated and suffers from biases introduced by manual inspection. Several parameters including window sizes to determine local expression levels, minimum coverage and significance thresholds to control for sensitivity and specificity have been fixed globally for this study.

We annotated 1421 putative TSSs in *Xcv* (Supplementary Table S2) including riboswitches and genes for conserved housekeeping and novel sRNAs. Interestingly, 178 TSSs correspond to antisense transcripts including six that map to type III effector genes and to *hrcC*, which is transcriptionally induced by HrpX and essential for T3S and pathogenicity (Supplementary Tables S2 and S4) (10,45). The potential role of post-transcriptional regulation in *Xcv* is further supported by the finding that 22% of all nucleotides that belong to annotated CDSs are covered by antisense reads. Nevertheless, the majority of these reads might be derived from promiscuous transcription initiation as it was also suggested for *E. coli* (46). It remains to be clarified whether the identified antisense transcripts in *Xcv* represent functional gene products or the transcription itself has a regulatory function.

We identified 831 putative primary TSSs, which were assigned to 17.35% of the 4726 annotated CDSs (Figure 1B and Supplementary Table S2) (9). Similarly, in the archeon *Methanosarcina mazei* TSSs for $\sim$20% of the CDSs were assigned (24). A considerably larger number of TSSs corresponding to 60% of the CDSs were recently mapped in the plant symbiont *Sinorhizobium meliloti* (31) and the human pathogen *H. pylori* ($>$50%) (4). This might be explained by the plethora of conditions analyzed and/or the higher number of sequencing reads and is supported by our finding that TSSs in *Xcv* are predominantly assigned to

CDSs with high expression level whereas CDSs without assigned TSS are generally weakly or not expressed (Supplementary Figure S2).

In *Xcv*, the majority of 5′-UTRs appears to be <50 bp (Figure 1D), which is characteristic for bacteria (3). Surprisingly, there is no clear consensus sequence for ribosome binding. A recent study (47) analyzed the evolutionary process of translation initiation in prokaryotes and found that a SD-initiated translation in xanthomonads is unlikely. In good agreement with this, we identified an unexpected high number of leaderless mRNAs in *Xcv* (Supplementary Table S5) suggesting an alternative mechanism of ribosome guidance. In *Xcv*, transcription of 82% of the leaderless mRNAs starts with AUG which was shown to be essential for stable ribosome binding to these transcripts in *E. coli* (48).

Unusually long 5′-UTRs as identified here for *Xcv* might be indicative of extensive post-transcriptional regulation, e.g. by sRNA-mediated modulation of mRNA translation or transcript stability. Since also 5′-UTRs of genes that encode type III effector proteins are unusually long (Supplementary Table S5), this might indicate a role of these 5′-UTRs in virulence. For instance, the genes for the type III effector proteins XopN and XopAA, shown to be important virulence factors of *Xcv* (49,50), comprise 5′-UTRs of 173 and 477 bp, respectively (Supplementary Table S5). In *H. pylori*, mRNAs of genes involved in pathogenesis also carry long 5′-UTRs (4).

Another potential implication of the high number of unusually long 5′-UTRs in *Xcv* is that the respective CDSs might be longer than predicted by the genome annotation as shown recently for the type III effector protein XopD (51). On the other hand, a number of CDSs are presumably shorter than annotated, because 71 internal TSSs are located within the first 50 bp of annotated CDSs (Supplementary Table S3). We also identified 12 expressed new loci with potential coding capacity (Supplementary Table S7) exemplified by *sX6* that encodes an 80 amino acid protein (Figure 3). Hence, this study contributes to a first refinement of CDS annotation in *Xcv*.

sRNAs represent important post-transcriptional regulators involved in a variety of processes such as quorum sensing (52) and virulence (53). In this study, the combination of manual and automatic inspection of the cDNA sequencing data and northern blots verified 23 sRNAs in *Xcv*, seven of which represent antisense RNAs (Table 1). For six of the antisense RNAs we also detected expression of the complementary mRNAs. It should be noted, however, that our data do not allow distinguishing between cells that express both transcripts at the same time and cells that either express the mRNA or the antisense RNA.

Notably, expression of five intergenic sRNAs and three antisense RNAs verified in this study was affected by the master regulators of *Xcv* virulence, HrpG and/or HrpX (Table 1) (16,18,19). Coregulation of sRNA expression with the T3S system clearly suggests a role of these transcripts in the interaction of *Xcv* with its host plant. As a proof-of-principle, we have demonstrated that *sX12* contributes to virulence of *Xcv* (Figure 4B). Lack of *sX12* does not affect bacterial growth inside the host and T3S, i.e. bacterial fitness is not impaired (Supplementary Figure S4). What might be the targets of sX12? Preliminary experiments did not reveal an effect of the absence of sX12 on selected *hrp* (T3S) genes, i.e. transcript and protein accumulation was unaltered. Instead of regulating mRNA targets, sX12 might control gene expression in a different manner, e.g. by binding to proteins, DNA or metabolites. Furthermore, *sX12* might impinge on the efficiency of the T3S system, similar to the *Salmonella typhimurium* sRNA IsrJ which accumulates under infection conditions. IsrJ positively contributes to invasion and effector translocation (54).

After our analysis was complete, the identification of eight sRNAs in *Xanthomonas oryzae* pv. *oryzae* (*Xoo*) strain PXO99A was reported (55). In agreement with our data, the *Xoo* sRNAs, *Xoo*3, *Xoo*4 and *Xoo*6, represent orthologs of the *Xcv* RNAs sX14, asX4 and sX1, respectively (Table 1). Contrary to *Xoo*4 (55), which is 145 nt, our analyses revealed that asX4 in *Xcv* is 309-nt long and encoded antisense to an annotated CDS. We also identified potential TSSs for the *Xcv* homologs of *Xoo*1 and *Xoo*5, whereas *Xcv* lacks homologs of described bacterial sRNA genes except for housekeeping RNAs. Vice versa, the majority of sRNAs identified in *Xcv* is restricted to the genera *Xanthomonas*, *Xylella* and *Stenotrophomonas* (Table 1) and thus, reflects the current taxonomy (56). An estimation of the total number of sRNAs in *Xcv* is hampered by the relatively small number of sequence reads and the fact that, for example, TSSs of sRNA genes in the proximity (≤300 bp) of downstream CDSs are classified as primary TSSs (see *sX1*; Table 1).

A remarkable finding of this study is the indication of frequent processing of *Xcv* sRNAs, which appears to be growth-phase dependent. In several studies, sRNA processing was shown to affect sRNA activity, e.g. GlmZ from *E. coli* which is cleaved and thus inactivated (57,58). The *E. coli* sRNA IstR-1 is rendered inactive by RNase III-dependent cleavage upon sRNA-mRNA interaction (59). In contrast, MicX from *Vibrio cholerae* is stabilized by RNaseE-mediated cleavage which does not impair its interaction with target-mRNAs (60). Which ribonucleases are involved in processing of the *Xcv* sRNAs is not known.

The analysis of additional knock-out mutants is needed to assess sRNA functions in *Xcv*. In case of virulence phenotypes, a challenge will be the identification of the targets. Besides possible effects of sRNAs on mRNAs the target can also be an RNA-binding protein. To the best of our knowledge, the only reported sRNAs involved in the regulation of virulence gene expression in plant pathogenic bacteria are members of the RsmB family which was studied in *Erwinia carotovora* ssp. *carotovora*. RsmB antagonizes the RNA-binding protein RsmA that acts as translational repressor (61–63). Although a major virulence function was reported for RsmA from *X. campestris* pv. *campestris* the interacting sRNAs are not known yet (5). The latter is complicated by the lack of CsrB/RsmB sequence homologs in xanthomonads.

SUPPLEMENTARY DATA

Supplementary Data are available at NAR Online: Supporting Information (SI), Supplementary Figures S1–4, Supplementary Tables S1–9 and Supplementary References [64–76].

ACKNOWLEDGEMENTS

FUNDING

REFERENCES

1. van Vliet,A.H. (2010) Next generation sequencing of microbial transcriptomes: challenges and opportunities. *FEMS Microbiol. Lett.*, **302**, 1–7.
2. Croucher,N.J. and Thomson,N.R. (2010) Studying bacterial transcriptomes using RNA-seq. *Curr. Opin. Microbiol.*, **13**, 619–624.
3. Sorek,R. and Cossart,P. (2010) Prokaryotic transcriptomics: a new view on regulation, physiology and pathogenicity. *Nat. Rev. Genet.*, **11**, 9–16.
4. Sharma,C.M., Hoffmann,S., Darfeuille,F., Reignier,J., Findeiss,S., Sittka,A., Chabas,S., Reiche,K., Hackermüller,J., Reinhardt,R. *et al.* (2010) The primary transcriptome of the major human pathogen *Helicobacter pylori*. *Nature*, **464**, 250–255.
5. Chao,N.X., Wei,K., Chen,Q., Meng,Q.L., Tang,D.J., He,Y.Q., Lu,G.T., Jiang,B.L., Liang,X.X., Feng,J.X. *et al.* (2008) The *rsmA*-like gene *rsmA_{Xcc}* of *Xanthomonas campestris* pv. *campestris* is involved in the control of various cellular processes, including pathogenesis. *Mol. Plant-Microbe Interact.*, **21**, 411–423.
6. Plener,L., Manfredi,P., Valls,M. and Genin,S. (2010) PrhG, a transcriptional regulator responding to growth conditions, is involved in the control of the type III secretion system regulon in *Ralstonia solanacearum*. *J. Bacteriol.*, **192**, 1011–1019.
7. Filiatrault,M.J., Stodghill,P.V., Bronstein,P.A., Moll,S., Lindeberg,M., Grills,G., Schweitzer,P., Wang,W., Schroth,G.P., Luo,S. *et al.* (2010) Transcriptome analysis of *Pseudomonas syringae* identifies new genes, noncoding RNAs, and antisense activity. *J. Bacteriol.*, **192**, 2359–2372.
8. Jones,J.B., Stall,R.E. and Bouzar,H. (1998) Diversity among xanthomonads pathogenic on pepper and tomato. *Annu. Rev. Phytopathol.*, **36**, 41–58.
9. Thieme,F., Koebnik,R., Bekel,T., Berger,C., Boch,J., Büttner,D., Caldana,C., Gaigalat,L., Goesmann,A., Kay,S. *et al.* (2005) Insights into genome plasticity and pathogenicity of the plant pathogenic bacterium *Xanthomonas campestris* pv. vesicatoria revealed by the complete genome sequence. *J. Bacteriol.*, **187**, 7254–7266.
10. Bonas,U., Schulte,R., Fenselau,S., Minsavage,G.V., Staskawicz,B.J. and Stall,R.E. (1991) Isolation of a gene-cluster from *Xanthomonas-campestris* pv. *vesicatoria* that determines pathogenicity and the hypersensitive response on pepper and tomato. *Mol. Plant-Microbe Interact.*, **4**, 81–88.
11. Szczesny,R., Büttner,D., Escolar,L., Schulze,S., Seiferth,A. and Bonas,U. (2010) Suppression of the AvrBs1-specific hypersensitive response by the YopJ effector homolog AvrBsT from *Xanthomonas* depends on a SNF1-related kinase. *New Phytol.*, **187**, 1058–1074.
12. Büttner,D. and Bonas,U. (2010) Regulation and secretion of *Xanthomonas* virulence factors. *FEMS Microbiol. Rev.*, **34**, 107–133.
13. White,F.F., Potnis,N., Jones,J.B. and Koebnik,R. (2009) The type III effectors of *Xanthomonas*. *Mol. Plant Pathol.*, **10**, 749–766.
14. White,F.F., Yang,B. and Johnson,L.B. (2000) Prospects for understanding avirulence gene function. *Curr. Opin. Plant Biol.*, **3**, 291–298.
15. Klement,Z. (1982) In: Mount,M.S. and Lacy,G.H. (eds), *Phytopathogenic Prokaryotes*, Vol. 2. Academic Press, New York, pp. 149–177.
16. Wengelnik,K. and Bonas,U. (1996) HrpXv, an AraC-type regulator, activates expression of five of the six loci in the *hrp* cluster of *Xanthomonas campestris* pv. vesicatoria. *J. Bacteriol.*, **178**, 3462–3469.
17. Schulte,R. and Bonas,U. (1992) Expression of the *Xanthomonas campestris* pv. vesicatoria *hrp* gene cluster, which determines pathogenicity and hypersensitivity on pepper and tomato, is plant inducible. *J. Bacteriol.*, **174**, 815–823.
18. Wengelnik,K., Van den Ackerveken,G. and Bonas,U. (1996) HrpG, a key *hrp* regulatory protein of *Xanthomonas campestris* pv. *vesicatoria* is homologous to two-component response regulators. *Mol. Plant-Microbe Interact.*, **9**, 704–712.
19. Noel,L., Thieme,F., Nennstiel,D. and Bonas,U. (2001) cDNA-AFLP analysis unravels a genome-wide *hrpG*-regulon in the plant pathogen *Xanthomonas campestris* pv. *vesicatoria*. *Mol. Microbiol.*, **41**, 1271–1281.
20. Koebnik,R., Krüger,A., Thieme,F., Urban,A. and Bonas,U. (2006) Specific binding of the *Xanthomonas campestris* pv. vesicatoria AraC-type transcriptional activator HrpX to plant-inducible promoter boxes. *J. Bacteriol.*, **188**, 7652–7660.
21. Wengelnik,K., Rossier,O. and Bonas,U. (1999) Mutations in the regulatory gene *hrpG* of *Xanthomonas campestris* pv. vesicatoria result in constitutive expression of all *hrp* genes. *J. Bacteriol.*, **181**, 6828–6831.
22. Hartmann,R.K., Bindereif,A., Schön,A. and Westhof,E. (2005) *Handbook of RNA biochemistry*, Vol. 2. Wiley-VCH, Weinheim, Germany, pp. 636–637.
23. Argaman,L., Hershberg,R., Vogel,J., Bejerano,G., Wagner,E.G., Margalit,H. and Altuvia,S. (2001) Novel small RNA-encoding genes in the intergenic regions of *Escherichia coli*. *Curr. Biol.*, **11**, 941–950.
24. Jäger,D., Sharma,C.M., Thomsen,J., Ehlers,C., Vogel,J. and Schmitz,R.A. (2009) Deep sequencing analysis of the *Methanosarcina mazei* Gö1 transcriptome in response to nitrogen availability. *Proc. Natl Acad. Sci. USA*, **106**, 21878–21882.
25. Skellam,J.G. (1946) The frequency distribution of the difference between two Poisson variates belonging to different populations. *J. R. Stat. Soc. Ser. A*, **109**, 296.
26. Fawcett,T. (2006) An introduction to ROC analysis. *Pattern Recogn. Lett.*, **27**, 861–874.
27. Hoffmann,S., Otto,C., Kurtz,S., Sharma,C.M., Khaitovich,P., Vogel,J., Stadler,P.F. and Hackermüller,J. (2009) Fast mapping of short sequences with mismatches, insertions and deletions using index structures. *PLoS Comput. Biol.*, **5**, 10.1371/journal.pcbi.1000502.
28. Albrecht,M., Sharma,C.M., Reinhardt,R., Vogel,J. and Rudel,T. (2010) Deep sequencing-based discovery of the *Chlamydia trachomatis* transcriptome. *Nucleic Acids Res.*, **38**, 868–877.

29. Wurtzel,O., Sapra,R., Chen,F., Zhu,Y., Simmons,B.A. and Sorek,R. (2010) A single-base resolution map of an archaeal transcriptome. *Genome Res.*, **20**, 133–141.

30. Mitschke,J., Georg,J., Scholz,I., Sharma,C.M., Dienst,D., Bantscheff,J., Voss,B., Steglich,C., Wilde,A., Vogel,J. *et al.* (2011) An experimentally anchored map of transcriptional start sites in the model cyanobacterium *Synechocystis* sp. PCC6803. *Proc. Natl Acad. Sci. USA*, **108**, 2124–2129.

31. Schlüter,J.P., Reinkensmeier,J., Daschkey,S., Evguenieva-Hackenberg,E., Janssen,S., Jänicke,S., Becker,J.D., Giegerich,R. and Becker,A. (2010) A genome-wide survey of sRNAs in the symbiotic nitrogen-fixing alpha-proteobacterium *Sinorhizobium meliloti*. *BMC Genomics*, **11**, 1111–1245.

32. Fenselau,S. and Bonas,U. (1995) Sequence and expression analysis of the *hrpB* pathogenicity operon of *Xanthomonas campestris* pv. *vesicatoria* which encodes eight proteins with similarity to components of the Hrp, Ysc, Spa, and Fli secretion systems. *Mol. Plant-Microbe Interact.*, **8**, 845–854.

33. Burgess,R.R. and Anthony,L. (2001) How sigma docks to RNA polymerase and what sigma does. *Curr. Opin. Microbiol.*, **4**, 126–131.

34. Findeiß,S., Schmidtke,C., Stadler,P.F. and Bonas,U. (2010) A novel family of plasmid-transferred anti-sense ncRNAs. *RNA Biol.*, **7**, 120–124.

35. Barrick,J.E., Corbino,K.A., Winkler,W.C., Nahvi,A., Mandal,M., Collins,J., Lee,M., Roth,A., Sudarsan,N., Jona,I. *et al.* (2004) New RNA motifs suggest an expanded scope for riboswitches in bacterial genetic control. *Proc. Natl Acad. Sci. USA*, **101**, 6421–6426.

36. Wang,J.X. and Breaker,R.R. (2008) Riboswitches that sense S-adenosylmethionine and S-adenosylhomocysteine. *Biochem. Cell Biol.*, **86**, 157–168.

37. Winkler,W.C. and Breaker,R.R. (2005) Regulation of bacterial gene expression by riboswitches. *Annu. Rev. Microbiol.*, **59**, 487–517.

38. Crossman,L.C., Gould,V.C., Dow,J.M., Vernikos,G.S., Okazaki,A., Sebaihia,M., Saunders,D., Arrowsmith,C., Carver,T., Peters,N. *et al.* (2008) The complete genome, comparative and functional analysis of *Stenotrophomonas maltophilia* reveals an organism heavily shielded by drug resistance determinants. *Genome Biol.*, **9**, R74.

39. da Silva,A.C., Ferro,J.A., Reinach,F.C., Farah,C.S., Furlan,L.R., Quaggio,R.B., Monteiro-Vitorello,C.B., Van Sluys,M.A., Almeida,N.F., Alves,L.M. *et al.* (2002) Comparison of the genomes of two *Xanthomonas* pathogens with differing host specificities. *Nature*, **417**, 459–463.

40. Remenant,B., Coupat-Goutaland,B., Guidot,A., Cellier,G., Wicker,E., Allen,C., Fegan,M., Pruvost,O., Elbaz,M., Calteau,A. *et al.* (2010) Genomes of three tomato pathogens within the *Ralstonia solanacearum* species complex reveal significant evolutionary divergence. *BMC Genomics*, **11**, 379.

41. El Yacoubi,B., Brunings,A.M., Yuan,Q., Shankar,S. and Gabriel,D.W. (2007) In planta horizontal transfer of a major pathogenicity effector gene. *Appl. Environ. Microbiol.*, **73**, 1612–1621.

42. Makarova,K.S., Wolf,Y.I. and Koonin,E.V. (2009) Comprehensive comparative-genomic analysis of type 2 toxin-antitoxin systems and related mobile stress response systems in prokaryotes. *Biol. Direct*, **4**, 10.1186/1745-6150-1184-1515.

43. Washietl,S., Findeiß,S., Müller,S.A., Kalkhof,S., von Bergen,M., Hofacker,I.L., Stadler,P.F. and Goldman,N. (2011) RNAcode: robust discrimination of coding and noncoding regions in comparative sequence data. *RNA*, **17**, 578–594.

44. Szczesny,R., Jordan,M., Schramm,C., Schulz,S., Cogez,V., Bonas,U. and Büttner,D. (2010) Functional characterization of the Xcs and Xps type II secretion systems from the plant pathogenic bacterium *Xanthomonas campestris* pv. *vesicatoria*. *New Phytol.*, **187**, 983–1002.

45. Wengelnik,K., Marie,C., Russel,M. and Bonas,U. (1996) Expression and localization of HrpA1, a protein of *Xanthomonas campestris* pv. vesicatoria essential for pathogenicity and induction ofthe hypersensitive reaction. *J. Bacteriol.*, **178**, 1061–1069.

46. Dornenburg,J.E., Devita,A.M., Palumbo,M.J. and Wade,J.T. (2010) Widespread antisense transcription in *Escherichia coli*. *MBio*, **46**, e00024–10.

47. Nakagawa,S., Niimura,Y., Miura,K. and Gojobori,T. (2010) Dynamic evolution of translation initiation mechanisms in prokaryotes. *Proc. Natl Acad. Sci. USA*, **107**, 6382–6387.

48. Brock,J.E., Pourshahian,S., Giliberti,J., Limbach,P.A. and Janssen,G.R. (2008) Ribosomes bind leaderless mRNA in *Escherichia coli* through recognition of their 5′-terminal AUG. *RNA*, **14**, 2159–2169.

49. Kim,J.G., Li,X., Roden,J.A., Taylor,K.W., Aakre,C.D., Su,B., Lalonde,S., Kirik,A., Chen,Y., Baranage,G. *et al.* (2009) *Xanthomonas* T3S Effector XopN Suppresses PAMP-Triggered Immunity and Interacts with a Tomato Atypical Receptor-Like Kinase and TFT1. *Plant Cell*, **21**, 1305–1323.

50. Morales,C.Q., Posada,J., Macneale,E., Franklin,D., Rivas,I., Bravo,M., Minsavage,J., Stall,R.E. and Whalen,M.C. (2005) Functional analysis of the early chlorosis factor gene. *Mol. Plant-Microbe Interact.*, **18**, 477–486.

51. Canonne,J., Marino,D., Noel,L.D., Arechaga,I., Pichereaux,C., Rossignol,M., Roby,D. and Rivas,S. (2010) Detection and functional characterization of a 215 amino acid N-terminal extension in the *xanthomonas* type III effector XopD. *PLoS One*, **5**, e15773.

52. Bejerano-Sagie,M. and Xavier,K.B. (2007) The role of small RNAs in quorum sensing. *Curr. Opin. Microbiol.*, **10**, 189–198.

53. Papenfort,K. and Vogel,J. (2010) Regulatory RNA in bacterial pathogens. *Cell Host Microbe*, **8**, 116–127.

54. Padalon-Brauch,G., Hershberg,R., Elgrably-Weiss,M., Baruch,K., Rosenshine,I., Margalit,H. and Altuvia,S. (2008) Small RNAs encoded within genetic islands of *Salmonella typhimurium* show host-induced expression and role in virulence. *Nucleic Acids Res.*, **36**, 1913–1927.

55. Liang,H., Zhao,Y.T., Zhang,J.Q., Wang,X.J., Fang,R.X. and Jia,Y.T. (2011) Identification and functional characterization of small non-coding RNAs in *Xanthomonas oryzae* pathovar *oryzae*. *BMC Genomics*, **12**, 101186/1471-2164-1112-1187.

56. Cutino-Jimenez,A.M., Martins-Pinheiro,M., Lima,W.C., Martin-Tornet,A., Morales,O.G. and Menck,C.F. (2010) Evolutionary placement of Xanthomonadales based on conserved protein signature sequences. *Mol. Phylogenet. Evol.*, **54**, 524–534.

57. Kalamorz,F., Reichenbach,B., Marz,W., Rak,B. and Görke,B. (2007) Feedback control of glucosamine-6-phosphate synthase GlmS expression depends on the small RNA GlmZ and involves the novel protein YhbJ in *Escherichia coli*. *Mol. Microbiol.*, **65**, 1518–1533.

58. Urban,J.H. and Vogel,J. (2008) Two seemingly homologous noncoding RNAs act hierarchically to activate *glmS* mRNA translation. *PLoS Biol.*, **6**, e64.

59. Vogel,J., Argaman,L., Wagner,E.G. and Altuvia,S. (2004) The small RNA IstR inhibits synthesis of an SOS-induced toxic peptide. *Curr. Biol.*, **14**, 2271–2276.

60. Davis,B.M. and Waldor,M.K. (2007) RNase E-dependent processing stabilizes MicX, a *Vibrio cholerae* sRNA. *Mol. Microbiol.*, **65**, 373–385.

61. Cui,Y., Chatterjee,A., Liu,Y., Dumenyo,C.K. and Chatterjee,A.K. (1995) Identification of a global repressor gene, *rsmA*, of *Erwinia carotovora* subsp. carotovora that controls extracellular enzymes, N-(3-oxohexanoyl)-L-homoserine lactone, and pathogenicity in soft-rotting Erwinia spp. *J. Bacteriol.*, **177**, 5108–5115.

62. Liu,Y., Cui,Y., Mukherjee,A. and Chatterjee,A.K. (1998) Characterization of a novel RNA regulator of *Erwinia carotovora* ssp. *carotovora* that controls production of extracellular enzymes and secondary metabolites. *Mol. Microbiol.*, **29**, 219–234.

63. Cui,Y., Chatterjee,A., Yang,H. and Chatterjee,A.K. (2008) Regulatory network controlling extracellular proteins in *Erwinia carotovora* subsp. carotovora: FlhDC, the master regulator of flagellar genes, activates *rsmB* regulatory RNA production by affecting *gacA* and *hexA* (*lrhA*) expression. *J. Bacteriol.*, **190**, 4610–4623.

64. Berezikov,E., Thuemmler,F., van Laake,L.W., Kondova,I., Bontrop,R., Cuppen,E. and Plasterk,R.H. (2006) Diversity of microRNAs in human and chimpanzee brain. *Nat. Genet.*, **38**, 1375–1377.

65. Bailey,T.L. and Elkan,C. (1995) The value of prior knowledge in discovering motifs with MEME. *Proc. Int. Conf. Intell. Syst. Mol. Biol.*, **3**, 21–29.

66. Blanchette,M., Kent,W.J., Riemer,C., Elnitski,L., Smit,A.F., Roskin,K.M., Baertsch,R., Rosenbloom,K., Clawson,H., Green,E.D. *et al.* (2004) Aligning multiple genomic sequences with the threaded blockset aligner. *Genome Res.*, **14**, 708–715.

67. Hertel,J., de Jong,D., Marz,M., Rose,D., Tafer,H., Tanzer,A., Schierwater,B. and Stadler,P.F. (2009) Non-coding RNA annotation of the genome of *Trichoplax adhaerens*. *Nucleic Acids Res.*, **37**, 1602–1615.

68. Will,S., Reiche,K., Hofacker,I.L., Stadler,P.F. and Backofen,R. (2007) Inferring noncoding RNA families and classes by means of genome-scale structure-based clustering. *PLoS Comput. Biol.*, **3**, e65.

69. Büttner,D., Nennstiel,D., Klüsener,B. and Bonas,U. (2002) Functional analysis of HrpF, a putative type III translocon protein from *Xanthomonas campestris* pv. vesicatoria. *J. Bacteriol.*, **184**, 2389–2398.

70. Rossier,O., Van den Ackerveken,G. and Bonas,U. (2000) HrpB2 and HrpF from *Xanthomonas* are type III-secreted proteins and essential for pathogenicity and recognition by the host plant. *Mol. Microbiol.*, **38**, 828–838.

71. Canteros,B.I. (1990), Ph.D. thesis. *University of Florida, Gainesville, FL.*

72. Menard,R., Sansonetti,P.J. and Parsot,C. (1993) Nonpolar mutagenesis of the *ipa* genes defines IpaB, IpaC, and IpaD as effectors of *Shigella flexneri* entry into epithelial cells. *J. Bacteriol.*, **175**, 5899–5906.

73. Bonas,U., Stall,R.E. and Staskawicz,B. (1989) Genetic and structural characterization of the avirulence gene *avrBs3* from *Xanthomonas campestris* pv. *vesicatoria. Mol. Gen. Genet.*, **218**, 127–136.

74. Huguet,E., Hahn,K., Wengelnik,K. and Bonas,U. (1998) *hpaA* mutants of *Xanthomonas campestris* pv. *vesicatoria* are affected in pathogenicity but retain the ability to induce host-specific hypersensitive reaction. *Mol. Microbiol.*, **29**, 1379–1390.

75. Figurski,D.H. and Helinski,D.R. (1979) Replication of an origin-containing derivative of plasmid RK2 dependent on a plasmid function provided *in trans. Proc. Natl Acad. Sci. USA*, **76**, 1648–1652.

76. Gardner,P.P., Daub,J., Tate,J., Moore,B.L., Osuch,I.H., Griffiths-Jones,S., Finn,R.D., Nawrocki,E.P., Kolbe,D.L., Eddy,S.R. *et al.* (2010) Rfam: Wikipedia, clans and the 'decimal' release. *Nucleic Acids Res.*, **39**, D141–145.

2.1.1.1. Anlagen zu Publikation 1

Die folgenden ,Supplementary Data' enthalten Zusatzinformationen zu Kapitel 2.1.1.: ,Supporting Information', Abbildung S1 bis S4 und Referenzen. Die Tabellen S1 bis S9 sind im Anhang aufgeführt.

Supporting Information

SI Materials and Methods

RNA isolation for 454 pyrosequencing, Northern blot and RACE analysis.
RNA was extracted for 454 pyrosequencing as follows: *Xcv* strains 85-10 and 85* were grown in NYG medium to exponential growth phase (OD$_{600}$ = 0.6). Then, 10 ml stop-solution (95% ethanol, 5% phenol) was added to 40 ml bacterial culture which was snap-frozen in liquid nitrogen, thawn on ice and centrifuged. Cells were resuspended in 6 ml buffer (0.02 M sodium acetate pH 5.5, 0.5% SDS, 1 mM EDTA). RNA was isolated by addition of 6 ml phenol, preheated to 60°C, followed by two chloroform extractions. The RNA was precipitated at -80°C overnight with 2.1 volumes of an ethanol/0.15 M sodium acetate solution. After centrifugation, the RNA was washed with 70% ethanol, dried, resuspended in water and treated with DNase I (Roche) followed by phenol-chloroform extraction.
For RACE and Northern blot analyses, RNA was isolated from NYG-grown *Xcv* strains at exponential and both exponential and stationary (OD$_{600}$ = ~1.5) growth phase, respectively, and treated with DNase I (Roche) as described (1).

5' and 3' RACE analyses.
RACE analyses (see Table 1) were carried out as described (2) with the following modifications: Reverse transcription was performed with 2 µg RNA, the Thermoscript RT system (Invitrogen) and a gene-specific primer for 5' RACE and a primer complementary to the 3' RNA adapter for 3' RACE analysis. Oligonucleotides used for RACE analyses are listed in Table S1. RACE-PCR was performed with Hotstar *Taq*-Polymerase (Qiagen). Cycling conditions: 95°C/15 min; 35 cycles of 95°C/40ss, 58°C/40 s, 72°C/40 s; 72°C/7 min. PCR products were cloned into pCR2.1-TOPO and transformed into *E.* coli TOP10F' (Invitrogen). Bacterial colonies were screened by colony PCR with vector-specific primers (see Table S1). Plasmid DNA was sequenced with an ABI PRISM 377 "Genetic Analyzer" DNA sequencer (Applied Biosystems).

Construction of cDNA libraries for dRNA-seq and 454 pyrosequencing.
Equal amounts of RNA from *Xcv* strains 85-10 and 85* were mixed. Next, we constructed dRNA-seq libraries as described (3). Briefly, primary transcripts of total RNA were enriched by a selective degradation of RNAs containing a 5' mono-phosphate (5'P) by treatment with Terminator^{TM} 5'P-dependent exonuclease (Epicentre). Prior to cDNA library construction, equal amounts of *Xcv* RNA were incubated 60 min at 30°C with terminator exonuclease (for generation of cDNA-library 2) or in buffer (for generation of cDNA-library 1). We used 1 unit terminator exonuclease per µg total RNA. Following organic extraction (25:24:1 v/v phenol/chloroform/isoamylalcohol), RNA was precipitated overnight with 2.5 volumes of an ethanol/0.1 M sodium acetate (pH 6.5) solution, and treated with 1 unit TAP (tobacco acid pyrophosphatase) (Epicentre) for 1 hour at 37°C to generate 5' mono-phosphates for linker ligation, and again purified by organic extraction and precipitation as above.
cDNA libraries for 454 pyrosequencing were constructed by *vertis* Biotechnology AG, Germany (http://www.vertisbiotech.com/) as described for eukaryotic microRNA (4) but omitting RNA size-fractionation prior to cDNA synthesis. Briefly, equal amounts of RNA treated with terminator exonuclease and untreated RNA, respectively, were poly(A)-tailed using poly(A) polymerase, followed by ligation of an RNA adapter to the 5´ P-RNA fragments. First-strand cDNA synthesis was performed using an oligo(dT)-adapter primer and M-MLV RNase H⁻ reverse transcriptase. Incubation temperatures were 42°C for 20 min, ramp to 55°C, followed by 55°C for 5 min. The cDNAs were PCR-amplified to yield a concentration of 20-30 ng/µl using a high fidelity DNA polymerase. Libraries were generated for the 454 FLX and Titanium kits. Each library contains a specific barcode sequence, which is attached to the 5' end of the cDNAs during PCR amplification: For FLX libraries, CCGA and CGCA were used as barcode tags for library 1 and 2, respectively. For Titanium libraries, ACGTGC and AGCGTA were used as barcode tags for library 1 and 2, respectively. 454 pyrosequencing was performed on a Roche 454 sequencer using FLX and Titanium chemistry at the Max Planck Institute for Molecular Genetics (Berlin, Germany). For library 1, a total of 62,056 and 98,293 reads was

sequenced using the FLX and Titanium kits, respectively. For library 2, a total of 51,091 and 98,505 reads was sequenced using the FLX and Titanium kits, respectively.

Sequence mapping.

For mapping of 454 reads, 5'-end-linker sequences were clipped, and reads with a poly(A) content of > 70% were discarded to prevent mapping errors. The remaining reads, including poly(A) tails and the 3' adapter sequence, were aligned to the genome sequence of *Xcv* strain 85-10 using the segemehl program (parameter settings E 10 A 65 D 1 H 2) (5). Mapped reads were post-processed by clipping of poorly aligned 3' ends as follows: For each alignment generated by segemehl (scoring scheme: match = 2, substitution = -2 and insertion/deletion = -3) the alignment score from the start of the read to each downstream nucleotide was calculated and stored in an array. All elements stored at a position greater than the maximum score at index i_{max} presumably correspond to the poly(A) tail and the 3'-linker sequence. Hence, the mapped read was clipped at position i_{max}. Reads that mapped with an identity of ≥ 85% and a minimum length of 12 nt were analyzed further whereas reads mapping to rRNA or tRNA genes were excluded.

Prediction of regulatory motifs and small ORFs.

Promoter regions, 50 nt upstream of the annotated TSS, and 5' UTRs were scanned with MEME (6) for regulatory motifs. To identify short conserved protein coding genes in *Xcv*, a multiple sequence alignment of 19 bacterial genomes (see Table S8) was calculated with the Multiz package (7). The alignments were analyzed for potential coding segments using RNAcode (8) and a p-value cutoff of 0.05. High scoring segments were combined if they were ≤ 15 nt apart and in the same reading frame. Regions that overlapped with annotated genes were discarded. The remaining 265 regions were inspected for potential open reading frames starting with an ATG and ending with a canonical stop codon. If no complete ORF was detected, the RNAcode high scoring segment was extended by 51 nt up- and downstream followed by repeated analysis. The RNAcode prediction resulted in 24 potential short ORFs in *Xcv* (Table S7; annotation files are available at www.bioinf.uni-leipzig.de/publications/supplements/10-035).

Rfam scan.

The *Rfam* database version 10.0 was downloaded from ftp://ftp.sanger.ac.uk/pub/databases/Rfam/10.0/. To scan the *Xcv* genome for known noncoding RNAs the *Rfam* provided Perl script rfam_scan.pl with an e-value cutoff of 100 was used. Eight riboswitches (FMN, SAH, Glycine, SAM, Cobalamin, TPP, yybP-ykoY) and five RNAs (RNase P, SRP, tmRNA, 6S-RNA, RrT) were identified (see Table S6). Annotation files are available at http://www.bioinf.uni-leipzig.de/publications/supplements/10-035.

Homology analysis.

Homology searches were based on scans of the bacterial NCBI genome database (ftp://ftp.ncbi.nih.gov/genomes/Bacteria/; downloaded 08/02/2010). To identify homologs of *Xcv* sRNA genes, Gotohscan (9) was used. Results were aligned with RNAclust, which is based on the LocARNA algorithm (10), and visualized with the SoupViewer (www.bioinf.uni-leipzig.de/software.html). Alignments of the analyzed *Xcv* sRNAs are available at http://www.bioinf.uni-leipzig.de/publications/supplements/10-035.

Protein detection.

The analysis of type III secretion was performed with *Xcv* strains incubated in minimal medium A as described (11). Total cell extracts and culture supernatants were concentrated 10 and 100 times, respectively, and were analyzed by sodium dodecyl sulfate–polyacrylamide gel electrophoresis (SDS-PAGE) and immunoblotting. For protein detection, specific polyclonal antibodies directed against HrpF (11), HrcJ (12) and GroEL (Stressgen) were used. A horseradish peroxidase-labeled anti-rabbit antibody (Amersham Pharmacia Biotech) was used as secondary antibody. The antibody reactions were visualized by enhanced chemiluminescence (Amersham Pharmacia Biotech).

For detection of the sX6-c-Myc protein, total cell extracts of NYG-grown bacteria (harvested at OD_{600} = 0.7) were concentrated 10-fold and analyzed by SDS-PAGE and immunoblotting using PVDF membranes. sX6-c-Myc was visualized with a monoclonal anti-c-Myc antibody (Roche) and a horseradish peroxidase-labeled anti-mouse secondary antibody (Amersham Pharmacia Biotech) by enhanced chemiluminescence (Amersham Pharmacia Biotech).

References
1. Hartmann, R.K., Bindereif, A., Schön, A. and Westhof, E. (2005) Handbook of RNA biochemistry. *Wiley-VCH, Weinheim, Germany*, **2**, 636-637.
2. Argaman, L., Hershberg, R., Vogel, J., Bejerano, G., Wagner, E.G., Margalit, H. and Altuvia, S. (2001) Novel small RNA-encoding genes in the intergenic regions of *Escherichia coli. Curr. Biol.*, **11**, 941-950.
3. Sharma, C.M., Hoffmann, S., Darfeuille, F., Reignier, J., Findeiss, S., Sittka, A., Chabas, S., Reiche, K., Hackermüller, J., Reinhardt, R. *et al.* (2010) The primary transcriptome of the major human pathogen *Helicobacter pylori*. *Nature*, **464**, 250-255.
4. Berezikov, E., Thuemmler, F., van Laake, L.W., Kondova, I., Bontrop, R., Cuppen, E. and Plasterk, R.H. (2006) Diversity of microRNAs in human and chimpanzee brain. *Nat. Genet.*, **38**, 1375-1377.
5. Hoffmann, S., Otto, C., Kurtz, S., Sharma, C.M., Khaitovich, P., Vogel, J., Stadler, P.F. and Hackermüller, J. (2009) Fast mapping of short sequences with mismatches, insertions and deletions using index structures. *PLoS Comput Biol*, **5**, 10.1371/journal.pcbi.1000502.
6. Bailey, T.L. and Elkan, C. (1995) The value of prior knowledge in discovering motifs with MEME. *Proc. Int. Conf. Intell. Syst. Mol. Biol.*, **3**, 21-29.
7. Blanchette, M., Kent, W.J., Riemer, C., Elnitski, L., Smit, A.F., Roskin, K.M., Baertsch, R., Rosenbloom, K., Clawson, H., Green, E.D. *et al.* (2004) Aligning multiple genomic sequences with the threaded blockset aligner. *Genome Res.*, **14**, 708-715.
8. Washietl, S., Findeiß, S., Müller, S.A., Kalkhof, S., von Bergen, M., Hofacker, I.L., Stadler, P.F. and Goldman, N. (2011) RNAcode: Robust discrimination of coding and noncoding regions in comparative sequence data. *RNA*, **17**, 578-594.
9. Hertel, J., de Jong, D., Marz, M., Rose, D., Tafer, H., Tanzer, A., Schierwater, B. and Stadler, P.F. (2009) Non-coding RNA annotation of the genome of *Trichoplax adhaerens*. *Nucleic Acids Res.*, **37**, 1602-1615.
10. Will, S., Reiche, K., Hofacker, I.L., Stadler, P.F. and Backofen, R. (2007) Inferring noncoding RNA families and classes by means of genome-scale structure-based clustering. *PLoS Comput. Biol.*, **3**, e65.
11. Büttner, D., Nennstiel, D., Klüsener, B. and Bonas, U. (2002) Functional analysis of HrpF, a putative type III translocon protein from *Xanthomonas campestris* pv. vesicatoria. *J. Bacteriol.*, **184**, 2389-2398.
12. Rossier, O., Van den Ackerveken, G. and Bonas, U. (2000) HrpB2 and HrpF from *Xanthomonas* are type III-secreted proteins and essential for pathogenicity and recognition by the host plant. *Mol. Microbiol.*, **38**, 828-838.

Figure S1
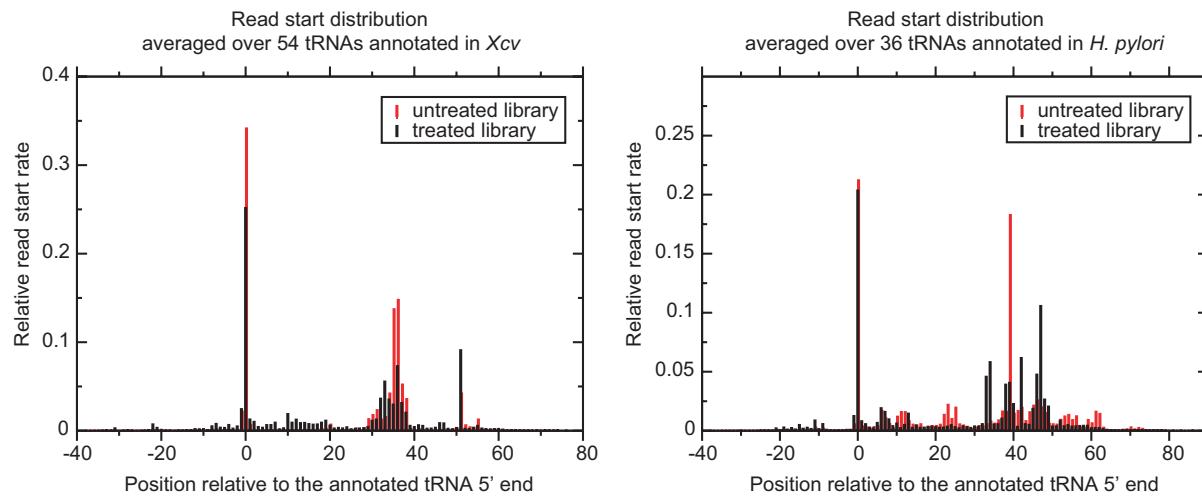
Figure S1. Averaged distribution of read starts across all annotated tRNA loci in *Xcv* and *H. pylori*, respectively.
For each library, the number of read starts at any position was normalized to the total number of reads that map to tRNAs. Position 0 indicates the 5' ends of annotated tRNA genes and corresponds to the RNase P processing site. Although treated libraries (black) show a relative reduction of read start rates compared to untreated RNA-seq libraries (red), tRNA expression is still detected in the enriched libraries. TSSs of tRNAs are expected to be located upstream of position 0.

Figure S2
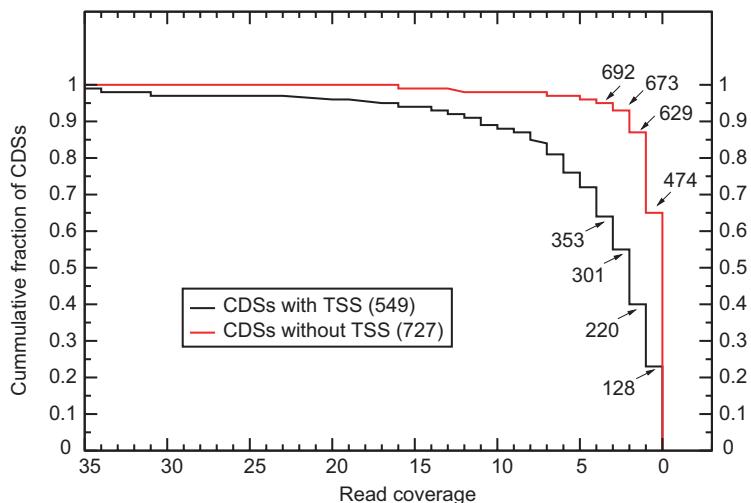
Figure S2. Coverage of CDSs with and without assigned TSSs in *Xcv*.
The plot displays the coverage of the first 100 nt of selected CDSs that are assumed to possess an own promoter since upstream genes are encoded on the opposite strand. The y-axis indicates the cummulative fraction of CDSs that exhibit a certain coverage in each data set. Numbers of CDSs with 0- up to 4-fold coverage are given at the corresponding positions within the plot. A successful TSS annotation depends on coverage. CDS without annotated TSS exhibit an overall lower coverage than CDSs with assigned TSS.

Figure S3 (Continued)
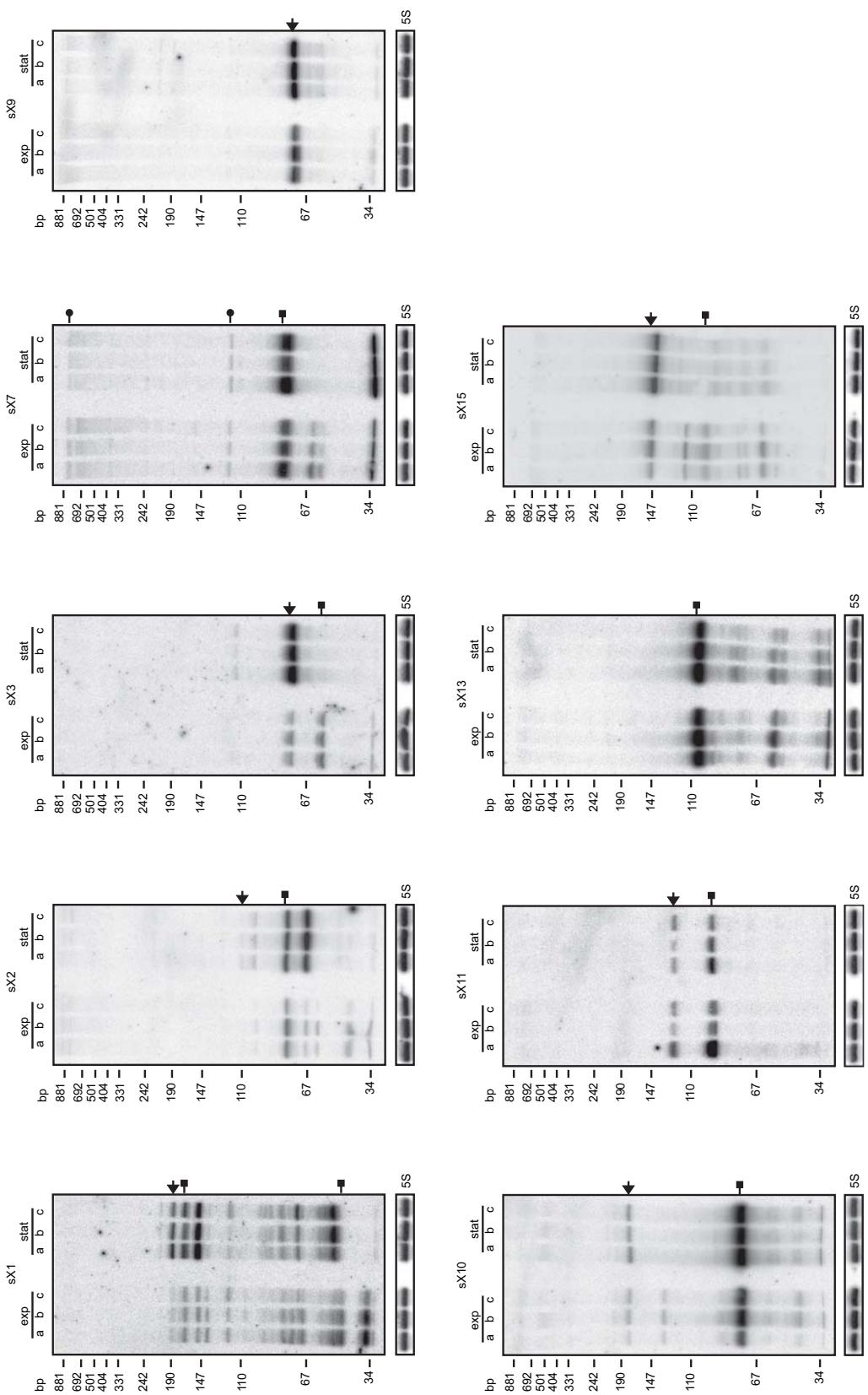
Figure S3
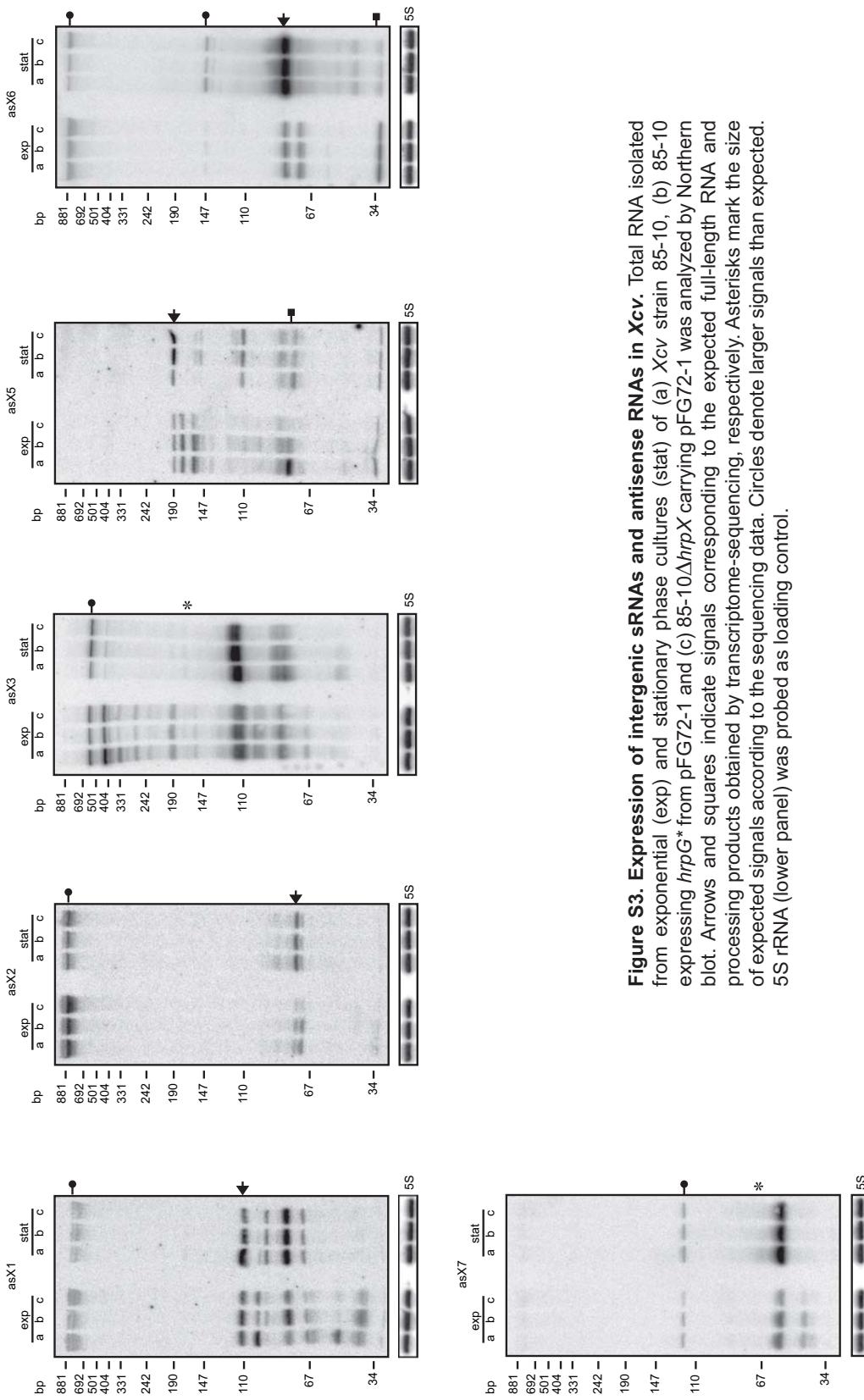
Figure S3. Expression of intergenic sRNAs and antisense RNAs in *Xcv*. Total RNA isolated from exponential (exp) and stationary phase cultures (stat) of (a) *Xcv* strain 85-10, (b) 85-10 expressing *hrpG** from pFG72-1 and (c) 85-10Δ*hrpX* carrying pFG72-1 was analyzed by Northern blot. Arrows and squares indicate signals corresponding to the expected full-length RNA and processing products obtained by transcriptome-sequencing, respectively. Asterisks mark the size of expected signals according to the sequencing data. Circles denote larger signals than expected. 5S rRNA (lower panel) was probed as loading control.

Figure S4

A

B
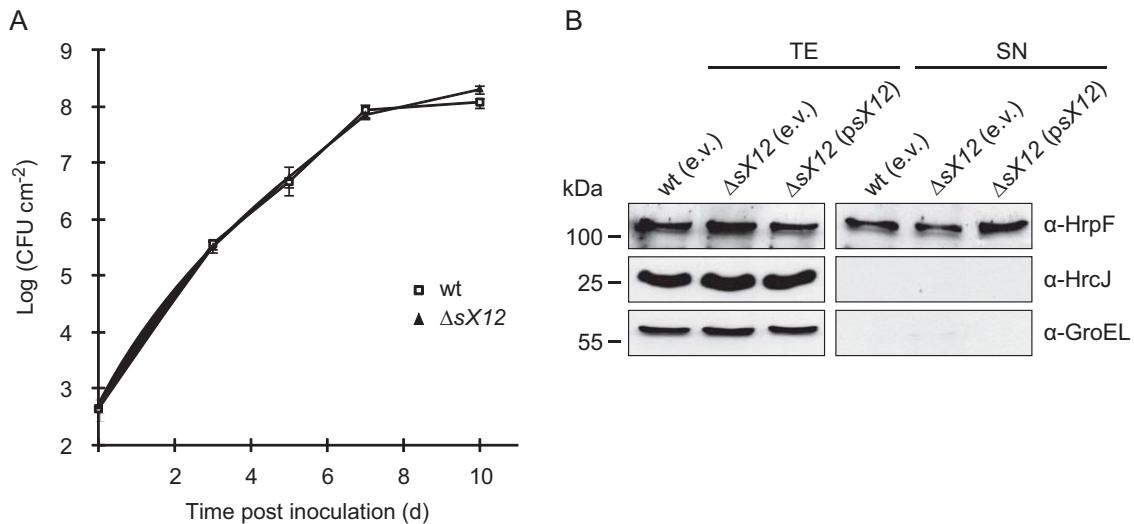
Figure S4. Deletion of *sX12* does not affect *in planta* growth and type III secretion. A. *In planta* growth of an *sX12* mutant strain. *Xcv* wild type strain 85-10 (wt) and an *sX12* deletion mutant (Δ*sX12*) were inoculated at a density of 10^4 CFU ml^{-1} into leaves of susceptible ECW pepper plants. Bacterial growth was determined over a period of 10 days. Data points indicate the mean of three samples from three different plants. Error bars represent standard deviations. B. Analysis of type III secretion. *Xcv* strain 85-10 carrying empty vector pLAFR6 [wt (e.v.)], an *sX12* deletion mutant [Δ*sX12* (e.v.)] and a complemented strain [Δ*sX12* (p*sX12*)] were incubated in secretion medium. The respective strains additionally express *hrpG** from pFG72-1. Total protein extracts (TE) and culture supernatants (SN) were analyzed by immunoblotting using antibodies directed against HrpF, HrcJ and GroEL.

2.1.1.2. Zusammenfassung der Ergebnisse

Der vorangegangene Artikel beschreibt die dRNA-Seq-basierte Identifizierung von TSSs und ncRNAs im *Xcv* Stamm 85-10. Die 454-Pyrosequenzierung ergab insgesamt 310.000 ‚reads', von denen 90.000, exklusive der ‚reads' für tRNAs und rRNAs, dem *Xcv* Genom zugewiesen werden konnten. Mittels eines neuartigen automatisierten Ansatzes zur Identifizierung von TSSs wurden 1.372 und 49 potentielle TSSs im *Xcv* Chromosom bzw. im Plasmid pXCV183 identifiziert. Die Klassifizierung der TSSs anhand ihrer Lokalisierung ergab 345 ‚antisense TSSs', welche auf dem Gegenstrang proteinkodierender Gene lokalisiert sind, sowie 426 ‚interne TSSs', die in annotierten ORFs lokalisiert sind. 831 der identifizierten TSSs repräsentieren vermutlich die TSSs von 17,35% der 4.726 in *Xcv* annotierten ORFs. Die korrespondierenden *Xcv* mRNAs enthalten keine konservierte SD-Sequenz und weisen überwiegend 5'-UTR Längen von 20-30 Nt auf. Dagegen weisen 14% und 13% der *Xcv* mRNAs keine bzw. ungewöhnlich lange 5'-UTRs (150-300 Nt) auf. Lange 5'-UTRs wurden auch für sechs mRNAs identifiziert, die Typ III Effektoren kodieren. Mittels bioinformatischer Analysen der *Xcv* Genomsequenz und der dRNA-Seq Daten wurden fünf in Bakterien konservierte RNAs mit vermutlich generellen zellulären Funktionen sowie acht potentielle Riboswitches identifiziert. Durch manuelle Sichtung der Sequenzierdaten und Northern Blot Analysen wurden 15 neue sRNAs, die konservierte 6S RNA sowie acht *cis*-kodierte asRNAs, einschließlich PtaRNA1 (s. Kapitel 2.2.1.), nachgewiesen. Zudem konnte gezeigt werden, dass die potentielle sRNA sX6 ein Protein kodiert, wohingegen die anderen sRNAs nicht-kodierend sind. Für die meisten sRNAs und asRNAs wurden in *Xcv* mögliche Prozessierungsprodukte detektiert, deren Abundanz abhängig von HrpG und/ oder HrpX oder der Wachstumsphase verändert war. Die Ko-Regulation der Expression bzw. Akkumulation von fünf sRNAs und drei asRNAs mit dem T3S System lässt eine Rolle in der Virulenz von *Xcv* vermuten und wurde am Beispiel der sRNA sX12 näher untersucht. Die Deletion des *sX12* Gens hatte eine verminderte Virulenz von *Xcv* in suszeptiblen und eine verzögerte HR Induktion in resistenten Pflanzen zur Folge, welche durch ektopische Expression von *sX12* komplementiert werden konnte. Das *in planta* Wachstum der *sX12* Deletionsmutante und *Xcv* 85-10 war vergleichbar. Zudem wurde nachgewiesen, dass die *in vitro* Typ III Sekretion des Translokonproteins HrpF bzw. die Akkumulation des HrcJ Proteins, einer zytoplasmatischen Komponente des T3S Systems, nicht durch die Deletion von *sX12* beeinflusst wird.

2.2. Bioinformatische Charakterisierung der *Xcv* asRNA PtaRNA1

2.2.1. Publikation 2

A novel family of plasmid-transferred anti-sense ncRNAs

Sven Findeiß,[1,]* Cornelius Schmidtke,[2] Peter F. Stadler[1,3-6] and Ulla Bonas[2]

[1]BioinformaticsGroup; Department of Computer Science; and InterdisciplinaryCenter for Bioinformatics; University of Leipzig; Leipzig, Germany;

[2]Institute of Biology; Department of Genetics; Martin-Luther-University Halle-Wittenberg; Halle, Germany; [3]Max-Planck-Institute for Mathematics in the Sciences; Leipzig, Germany; [4]Institute for Theoretical Chemistry; University of Vienna; Wien, Austria; [5]Fraunhofer Institute for Cell Therapy und Immunology; Leipzig, Germany; [6]Santa Fe Institute; Santa Fe, NM USA

The genome of Xanthomonas campestris pv. vesicatoria encodes a constitutively expressed small RNA, which we designate PtaRNA1, "Plasmid transferred anti-sense RNA". It exhibits all hallmarks of a novel RNA antitoxin that proliferates by frequent horizontal transfer. It shows an erratic phylogenetic distribution with occurrences on chromosomes in a few individual strains distributed across both beta- and gammaproteobacteria. Moreover, a homologous gene located on plasmid pMATVIM-7 of Pseudomonas aeruginosa is found. All ptaRNA1 homologs are located anti-sense to a putative toxin, which in turn is never encountered without the small RNA. The secondary structure of PtaRNA1, furthermore, is very similar to that of the FinP anti-sense RNA found on F-like plasmids in Escherichia coli.

Introduction

Several toxin-antitoxin systems of type 1, in which the toxin is a short protein and the antitoxin an anti-sense RNA and of type 2, where both elements are proteins, are frequently found in both prokaryotic chromosomes and plasmids.[1-4] The paradigmatic example for type 1 is the plasmid encoded *hok/sok* system in *Escherichia coli* and its close relatives. The toxin-encoding stable mRNA encodes a protein that rapidly leads to cell-death unless its translation is suppressed by a short-lived small RNA. The plasmid encoded module prevents the growth of plasmid-free offsprings thus ensuring the persistence of the plasmid in the population: After

cell division, plasmid-free cells still contain the stable toxin mRNA, while the comparably unstable antitoxin is quickly depleted. It is poorly understood how the chromosomally encoded systems function. Interestingly, the SOS-induced genes *tisB* and *symE* are expressed under very specific stress conditions. The corresponding antitoxins (SymR and Sib) are constitutively expressed.

Although distinct toxin-antitoxin systems have been found in widely separated bacterial groups (e.g., *hok/sok* in *E. coli* and *txpA/ratA* in *Bacillus subtilis*[5]), each of the known examples exhibits a very narrow phylogenetic distribution.

In this contribution we characterize by computational means a small RNA that has all the hallmarks of the known type 1 toxin-antitoxin systems but shows a rather wide spread erratic phylogenetic distribution that hints at frequent horizontal gene transfers.

Results

The founding member, PtaRNA1 ("Plasmid transferred anti-sense RNA"), of the family was detected in a library of pyrosequencing data of *Xanthomonas campestris* pv. *vesicatoria* strain 85-10 (*Xcv*) that was prepared and analyzed for unrelated purposes. The superposition of the individual reads revealed a small RNA encoded adjacent to the *trbL* gene. Expression and approximate size of the small RNA was verified by northern blot (**Fig. 1**). These analyses revealed a constitutive expression with respect to the tested growth phases. Interestingly,
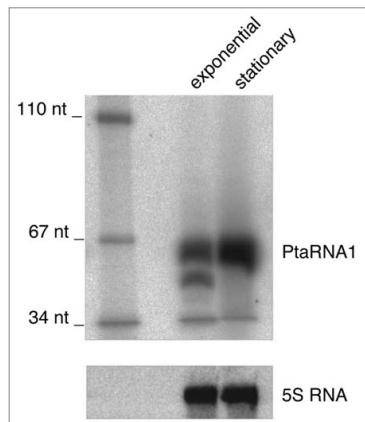
Figure 1. Expression of PtaRNA1 and 5S RNA in exponential and stationary growth phase of *Xcv* analysed by northern blot. The size of corresponding marker bands is indicated on the left.

two bands which indicate procession of the full length PtaRNA1 are detected in the exponential but not in the stationary growth phase.

Chromosomally encoded homologs of *ptaRNA1* were found in beta-proteobacteria (*Nitrosomonas eutropha* C91, *Azoarcus* sp. EbN1, *Verminephrobacter eiseniae* EF01-2, *Burkholderia cenocepacia* J2315, *B. pseudomallei* K96243, *B. pseudomallei* 9, *B. pseudomallei* 91, and *Acidovorax* JS42) as well as gamma-proteobacteria (*X. campestris* pv. *vasculorum* NCPPB702, *Shigella flexneri* 2a 2457T, *Acinetobacter baumannii* ATCC 17978, *Marinobacter aquaeolei* VT8, and *Pseudomonas aeruginosa* UCBPP-PA14). Two *ptaRNA1* copies were found in *N. eutropha* C91 and are named *ptaRNA1*-a and *ptaRNA1*-b. This observation is in accordance with various reported insertion, duplication

and rearrangement events within this species.[6]

Conspicuously, *ptaRNA1* was not found in other closely related genomes, e.g., other strains of Burkholderia, Pseudomonas or Xanthomonas. This distinguishes PtaRNA1 from most other bacterial small RNAs, such as the cyanobacteria-specific Yfr RNAs.[7] In addition to the chromosomal loci listed above, we found a *ptaRNA1* homolog in the *P. aeruginosa* plasmid pMATVIM-7, adjacent to the transfer region, a gene cluster that encodes proteins of unknown function and the plasmid stabilization protein ParE. **Figure 2A** depicts the alignment and the resulting consensus secondary structure of all detected PtaRNA1 homologs.

Phylogenetic analysis of the PtaRNA1 sequences (**Fig. 3**) shows that the
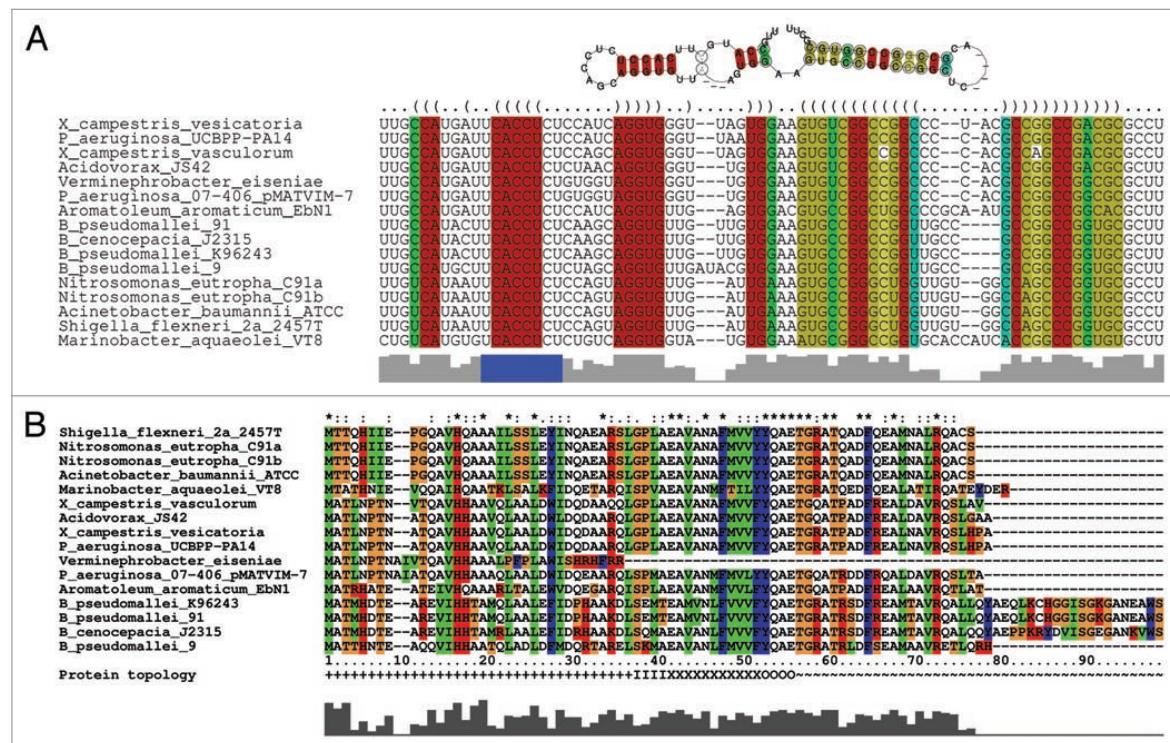
Figure 2. (A) Consensus secondary structure model of PtaRNA1 based on the depicted seed alignment. The structure is highly stable (minimum free energy -37 kcal/mol) and supported by various compensatory mutations within the stem on the right-hand side. Marked in blue is the region complementary to the putative Shine-Dalgarno sequence of the *XCV2162* mRNA. (B) Amino acid alignment of XCV2162 homologs. The alignment shows various totally (indicated by '*') and by substitutions (indicated by ':' and '.') supported and therefore conserved columns. The protein topology of a trans-membrane domain, predicted by MEMSAT3,[8] is indicated as well. '+' marks inside loop, '~' outside loop, 'O' outside helix cap, 'X' central transmembrane helix segment and 'I' inside cap. The truncated Verminephrobacter sequence was not used for the calculation of the conservation track.

Figure 3. Phylogenetic tree based on PtaRNA1 alignment (similar for XCV2162 alignment, data not shown). Class of the "host" species is shown by the symbols on the right hand side. Numbers indicate bootstrap values of the inner nodes.

only *ptaRNA1* encoding species in which no *trbL* homolog was found.

Analysis of the putative *ptaRNA1* promoter regions revealed the existence of two highly conserved sequence motifs of eight nucleotides. The first one starts between 42 and 36 nt upstream and the other 13/12 nt upstream of the transcription start corresponding to the -35 and -10 box, respectively (**Fig. 4**). In the upstream region of *XCV2162*, an ultra conserved AG-rich motif was found (**Fig. 4**), probably representing the Shine-Dalgarno sequence of the mRNA. This motif is entirely covered by the complementary *ptaRNA1* sequence (**Fig. 2A**).

The consensus secondary structure of PtaRNA1 consists of a 5'-stem loop and a long 3' stem which presumably acts as terminator hairpin, (**Fig. 2A**). This structure is very similar to that of FinP found in *E. coli* (data not shown). Interestingly, FinP is the anti-sense regulator of TraJ, a transcriptional activator required for expression of various conjungative protein components.[9] Thus, FinP is only one component in a complex network of several interacting molecules.

Discussion

All evidence available for PtaRNA1 suggests that *XCV2162*/*ptaRNA1* is a novel toxin-antitoxin system: both genes are only found as combined cluster and in fact do not appear as single genes; *XCV2162* encodes a relatively short protein that shows the typical topology of toxins with a trans-membrane domain; the presence on a plasmid in combination with the erratic phylogenetic distribution of the system indicates a frequent horizontal gene transfer. It is known that type 2 systems, where both molecules are proteins, show an erratic phylogenetic distribution.[4] We assume that this might also be the case for type 1 systems, such as the one presented here.

Furthermore, the phylogenetic distribution of the *XCV2162*/*ptaRNA1* pair indicates a very rapid loss of the chromosomal homologs: the erratic distribution suggests that we only see very recent chromosomal insertions. The homolog in Verminephrobacter with its truncated *XCV2162* coding sequence might

phylogeny of the PtaRNA1 sequences is not congruent with the phylogeny of their "host" species. This indicates that the proliferation of *ptaRNA1* depends on frequent horizontal transfer, presumably by means of plasmids.

The *ptaRNA1* gene in *Xcv* is located anti-sense to a so far uncharacterized protein coding gene (*XCV2162*). The gene is adjacent to *trbL*, encoding a type IV secretion system protein. The small overlap of both genes strongly suggests that PtaRNA1 is an anti-sense regulator of *XCV2162* (**Fig. 4**). We therefore searched the complete set of eubacterial genomes for homologs of *XCV2162* and found that *ptaRNA1* and *XCV2162* co-occur in all cases, indicating their functional linkage. This is in particular also the case in the *P. aeruginosa* plasmid pMATVIM-7,

whose uncharacterized gene *p07-406.22* is an ortholog of *XCV2162*, as in *Xcv* adjacent to *trbL* (**Fig. 4**). According to MEMSAT3[8] prediction, XCV2162 contains a trans-membrane domain (**Fig. 2B**), as it is also the case for many reported toxic proteins.[3,4]

Furthermore, the gene phylogeny of the XCV2162 proteins (not shown) is congruent with the phylogeny of PtaRNA1 sequences, indicating that they are transferred together. We observed a frequent co-occurrence of *ptaRNA1*/*XCV2162* and *trbL* homologs, albeit *trbL* was detected in at least 155 eubacterial genomes, suggesting that *trbL* might have a role in the frequent chromosomal insertions of the *ptaRNA1*/*XCV2162* system. In *V. eiseniae* EF01-2 we found a truncated *XCV2162* homolog. Verminephrobacter is also the

Figure 4. Surrounding genomic location of the *ptaRNA1* gene in *Xcv*. On the plus strand the coding sequences of *trbL* and *XCV2162* are indicated. In front of *XCV2162* an ultra conserved AG rich motif, the putative Shine-Dalgarno sequence is shown. *ptaRNA1* is encoded on the minus strand. A conserved -10 as well as -35 box (sequence logos) was found directly upstream of this gene.

represent the first step towards the complete loss of the system.

This apparent evolutionary instability further supports the hypothesis that *XCV2162/ptaRNA1* is a toxin-antitoxin pair. Only chromosomal integration of the toxin-antitoxin pair makes the plasmid dispensable. Thus, cell-death is prevented by chromosomal integration of the system. Plasmid-loss and subsequent destruction of the toxin *XCV2162* then leaves the chromosomal copy of the antitoxin PtaRNA1 without function, so it is also rapidly removed from the genome.

Materials and Methods

For northern blot analysis *Xanthomonas campestris* pv. *vesicatoria* strain 85-10 was cultivated at 30°C in nutrient-yeast-glycerol medium.[10] Cells were harvested at exponential and stationary growth phase at OD of 0.6 and 1.5, respectively. RNA was extracted as described in.[11] Northern blots were done according to[12] with the following modifications: 20 µg RNA were separated on 8.3 M urea—6% polyacrylamide gels. For detection of PtaRNA1 and 5S rRNA membranes were incubated for 1 h at 42°C with Rapid-hyb™ Buffer (GE Healthcare) containing ^{32}P 5' end-labeled oligodeoxyribonucleotides NB39 (5'-ATG GAG AGG TGA ATC ATG GC-3') and NB5S (5'-ATG ACC TAC TCT CGC ATG GC-3'), respectively.

Homology searches were based on scans of the bacterial NCBI genome database (ftp://ftp.ncbi.nih.gov/genomes/Bacteria/downloaded 06/12/2009) as well

as the plasmid genome database (http://www.genomics.ceh.ac.uk/plasmiddb/downloaded 06/12/2009). Homologs of protein coding genes were searched using tblastn of the Blast package.[13] Since non-coding RNAs may vary in sequence but still fold into the same secondary structure a semi-global alignment implementation, GotohScan,[14] was used to scan for *ptaRNA1* homologs. The microbial web Blast (http://www.ncbi.nlm.nih.gov/sutils/genom_table.cgi) was used to search for additional homologs especially in unfinished genome projects.

Alignments were calculated with ClustalW[15] and locARNATE[16] for sequence structure alignments, respectively. The consensus structure model was calculated with RNAalifold.[17]

Using, MEME[18] we analyzed the 100 nt upstream region of all homologous *ptaRNA1* loci. MEME searches for similarities among the given sequences and calculates descriptors for these motifs. To search for known regulatory sites within the 100 nt upstream region the PRODORIC database was queried using the Virtual Footprint v3.0 web tool.[19]

Submitted data. This manuscript documents the seed alignment ptaRNA1.seed.stk. A short summary is available as Wikipedia Entry at http://en.wikipedia.org/wiki/User:SveFinBioInf/ptaRNA1.

Acknowledgements

In order to keep the list of references at reasonable length, we had to give priority to reviews and recent publications.

Note

Supplementary materials can be found at:www.landesbioscience.com/supplement/FindeissRNA7-2-Sup.txt

References

1. Gerdes K, Wagner EGH. RNA antitoxin. Cur Op Microbiol 2007; 10:117-24.
2. Weaver KE. Emerging plasmid-encoded antisense RNA regulated systems. Curr Opin Microbiol 2007; 10:110-6.
3. Fozo EM, Hemm MR, Storz G. Small toxic proteins and the antisense RNAs that repress them. Microbiol Mol Biol Rev 2008; 72:579-89.
4. Makarova KS, Wolf YI, Koonin EV. Comprehensive comparative-genomic analysis of type 2 toxin-antitoxin systems and related mobile stress response systems in prokaryotes. Biol Direct 2009; 4:19.
5. Silvaggi JM, Perkins JB, Losick R. Small untranslated RNA antitoxin in *Bacillus subtilis*. J Bacteriol 2005; 187:6641-50.
6. Stein LY, et al. Whole-genome analysis of the ammonia-oxidizing bacterium, *Nitrosomonas eutropha* C91: implications for niche adaptation. Environ Microbiol 2007; 9:2993-3007.
7. Voss B, Gierga G, Axmann IM, Hess WR. A motif-based search in bacterial genomes identifies the ortholog of the small RNA Yfr1 in all lineages of cyanobacteria. BMC Genomics 2007; 8:375.
8. Nugent T, Jones DT. Transmembrane protein topology prediction using support vector machines. BMC Bioinformatics 2009; 10:159.
9. Arthur DC, Ghetu AF, Gubbins MJ, Edwards RA, Frost LS, Glover JNM. FinO is an RNA chaperone that facilitates sense-antisense RNA interactions. EMBO J 2003; 22:6346-55.
10. Daniels MJ, Barber CE, Turner PC, Sawczyc MK, Byrde RJW, Fielding AH. Cloning of genes involved in pathogenicity of *Xanthomonas campestris* pv. *campestris* using the broad host range cosmid pLAFR1. EMBO J 1984; 3:3323-8.
11. Hartmann RK, Bindereif ASA, Westhof E. Handbook of RNA Biochemistry. Wiley-VCH 2005.
12. Urban JH, Vogel J. Translational control and target recognition by *Escherichia coli* small RNAs in vivo. Nucleic Acids Res 2007; 35:1018-37.
13. Altschul SF, et al. Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. Nucleic Acids Res 1997; 25:3389-402.

14. Hertel J, et al. Non-coding RNA annotation of the genome of trichoplax adhaerens. Nucleic Acids Res 2009; 37:1602-15.

15. Larkin MA, et al. Clustal W and Clustal X version 2.0. Bioinformatics 2007; 23:2947-8.

16. Otto W, Will S, Backofen R. Structure local multiple alignment of RNA. In *Proceedings of the German Conference on Bioinformatics (CGB 08)*, volume P-136 of *LNI* GI 2008; 178-88.

17. Bernhart SH, Hofacker IL, Will S, Gruber AR, Stadler PF. RNAalifold: improved consensus structure prediction for RNA alignments. BMC Bioinformatics 2008; 9:474.

18. Bailey TL, Elkan C. The value of prior knowledge in discovering motifs with MEME. Proc Int Conf Intell Syst Mol Biol 1995; 3:21-9.

19. Münch R, et al. Virtual Footprint and PRODORIC: an integrative framework for regulon prediction in prokaryotes. Bioinformatics 2005; 21:4187-9.

2.2.1.1. Zusammenfassung der Ergebnisse

Der Artikel beschreibt die bioinformatische Charakterisierung der *Xcv* asRNA PtaRNA1 (‚plasmid-transferred antisense RNA 1‘), welche in der vorangegangenen dRNA-Seq Analyse (s. Kapitel 2.1.1.) als ncRNA Kandidat identifiziert wurde. Northern Blot Analysen von *Xcv* 85-10 ergaben, dass PtaRNA1 (72 Nt) in der stationären Wachstumsphase als stabiles Transkript akkumuliert, wohingegen die RNA in der exponentiellen Wachstumsphase vermutlich prozessiert wird. In *Xcv* überlappt das chromosomale *ptaRNA1* Gen in antisense Orientierung mit der 5‘-Region des *XCV2162* Gens, welches ein potentielles Transmembranprotein mit unbekannter Funktion kodiert. Zudem ist *ptaRNA1* benachbart zu *trbL* lokalisiert, welches vermutlich am konjugalen DNA Transfer beteiligt ist. Phylogenetische Analysen ergaben, dass Orthologe von *ptaRNA1* und *XCV2162* in den Chromosomen zahlreicher, nur entfernt verwandter ß- und γ-Proteobakterien konserviert sind und stets ko-lokalisieren, jedoch nicht in nahe verwandten Bakterien vorkommen. Zudem sind die entsprechenden Loci meist in Nachbarschaft von *trbL* lokalisiert. In Vertretern der Gattung *Xanthomonas* kommt *ptaRNA1* nur in *Xcv* und *X. campestris* pv. *vasculorum* vor. Der *ptaRNA1* Lokus weist typische Merkmale eines Typ I-Toxin-Antitoxin Systems auf, bei dem die Synthese eines toxischen Proteins durch eine *cis*-kodierte asRNA unterdrückt wird. Die sporadische phylogenetische Verbreitung des Lokus steht im scheinbaren Gegensatz zur hohen Sequenzkonservierung, deutet jedoch auf einen Erwerb durch horizontalen Gentransfer hin. Hierbei spielen vermutlich Plasmide eine Rolle, da ein entsprechender plasmidlokalisierter *ptaRNA1* Lokus in *P. aeruginosa* vorhergesagt wurde.

2.3. Funktionelle Charakterisierung der *Xcv* sRNA sX13

2.3.1. Publikation 3

Small RNA sX13: A Multifaceted Regulator of Virulence in the Plant Pathogen *Xanthomonas*

Cornelius Schmidtke[1]*, Ulrike Abendroth[1], Juliane Brock[1], Javier Serrania[2], Anke Becker[2], Ulla Bonas[1]*

1 Institute for Biology, Department of Genetics, Martin-Luther-Universität Halle-Wittenberg, Halle, Germany, **2** Loewe Center for Synthetic Microbiology and Department of Biology, Philipps-Universität Marburg, Marburg, Germany

Abstract

Small noncoding RNAs (sRNAs) are ubiquitous posttranscriptional regulators of gene expression. Using the model plant-pathogenic bacterium *Xanthomonas campestris* pv. *vesicatoria* (*Xcv*), we investigated the highly expressed and conserved sRNA sX13 in detail. Deletion of *sX13* impinged on *Xcv* virulence and the expression of genes encoding components and substrates of the Hrp type III secretion (T3S) system. qRT-PCR analyses revealed that sX13 promotes mRNA accumulation of HrpX, a key regulator of the T3S system, whereas the mRNA level of the master regulator HrpG was unaffected. Complementation studies suggest that sX13 acts upstream of HrpG. Microarray analyses identified 63 sX13-regulated genes, which are involved in signal transduction, motility, transcriptional and posttranscriptional regulation and virulence. Structure analyses of *in vitro* transcribed sX13 revealed a structure with three stable stems and three apical C-rich loops. A computational search for putative regulatory motifs revealed that sX13-repressed mRNAs predominantly harbor G-rich motifs in proximity of translation start sites. Mutation of sX13 loops differentially affected *Xcv* virulence and the mRNA abundance of putative targets. Using a GFP-based reporter system, we demonstrated that sX13-mediated repression of protein synthesis requires both the C-rich motifs in sX13 and G-rich motifs in potential target mRNAs. Although the RNA-binding protein Hfq was dispensable for sX13 activity, the *hfq* mRNA and Hfq::GFP abundance were negatively regulated by sX13. In addition, we found that G-rich motifs in sX13-repressed mRNAs can serve as translational enhancers and are located at the ribosome-binding site in 5% of all protein-coding *Xcv* genes. Our study revealed that sX13 represents a novel class of virulence regulators and provides insights into sRNA-mediated modulation of adaptive processes in the plant pathogen *Xanthomonas*.

Introduction

The survival and prosperity of bacteria depends on their ability to adapt to a variety of environmental cues such as nutrient availability, osmolarity and temperature. Besides the adaptation to the environment by transcriptional regulation of gene expression bacteria express regulatory RNAs that modulate expression on the posttranscriptional level [1,2]. Small regulatory RNAs (sRNAs; ∼50–300 nt) have been intensively studied in the enterobacteria *Escherichia coli* and *Salmonella* spp. and, in most cases, regulate translation and/or stability of target mRNAs through short and imperfect base-pairing (10 to 25 nucleotides) [1,3,4,5]. The majority of characterized sRNAs inhibits translation of target mRNAs by pairing near or at the ribosome-binding site (RBS) [1,6]. In addition, sRNAs can promote target mRNA translation, e. g., the sRNAs ArcZ, DsrA and RprA activate translation of sigma factor RpoS [7,8,9]. Regulation of multiple rather than single genes has emerged as a major feature of sRNAs affecting processes like iron homeostasis, carbon metabolism, stress responses and quorum sensing (QS) [1,2,6]. In numerous cases, sRNAs are under transcriptional control of two-component systems (TCS), which themselves are often controlled by sRNAs

[10]. The activity and stability of most enterobacterial sRNAs requires the hexameric RNA-binding protein Hfq, which facilitates the formation of sRNA-mRNA duplexes and their subsequent degradation by the RNA degradosome [1,11]. Hfq is present in approximately 50% of all bacterial species and acts in concert with sRNAs to regulate stress responses and virulence of a number of animal- and human-pathogenic bacteria [5,12].

To date, little is known about sRNAs in plant-pathogenic bacteria. Only recently, high throughput RNA-sequencing approaches uncovered potential sRNAs in the plant-pathogenic α-proteobacterium *Agrobacterium tumefaciens* [13], the γ-proteobacteria *Pseudomonas syringae* pv. *tomato* [14] and *Xanthomonas campestris* pv. *vesicatoria* (*Xcv*) [15,16]. Additional studies identified four and eight sRNAs in *X. campestris* pv. *campestris* (*Xcc*) and *X. oryzae* pv. *oryzae* (*Xoo*), respectively [17,18]. So far, only few sRNAs of plant-pathogenic bacteria were characterized with respect to potential targets. Examples include the *A. tumefaciens* antisense RNA RepE and the sRNA AbcR1, which regulate Ti-plasmid replication and the expression of ABC transporters, respectively [19,20]. RNAs involved in virulence of plant-pathogenic bacteria were so far only reported for *Erwinia* spp. and *Xcv*. In *Erwinia*, the protein-binding RNA RsmB modulates

Author Summary

Since the discovery of the first regulatory RNA in 1981, hundreds of small RNAs (sRNAs) have been identified in bacteria. Although sRNA-mediated control of virulence was demonstrated for numerous animal- and human-pathogenic bacteria, sRNAs and their functions in plant-pathogenic bacteria have been enigmatic. We discovered that the sRNA sX13 is a novel virulence regulator of *Xanthomonas campestris* pv. *vesicatoria* (*Xcv*), which causes bacterial spot disease on pepper and tomato. sX13 contributes to the *Xcv*-plant interaction by promoting the synthesis of an essential pathogenicity factor of *Xcv*, i. e., the type III secretion system. Thus, in addition to transcriptional regulation, sRNA-mediated posttranscriptional regulation contributes to virulence of plant-pathogenic xanthomonads. To repress target mRNAs carrying G-rich motifs, sX13 employs C-rich loops. Hence, sX13 exhibits striking structural similarity to sRNAs in distantly related human pathogens, e. g., *Staphylococcus aureus* and *Helicobacter pylori*, suggesting that structure-driven target regulation via C-rich motifs represents a conserved feature of sRNA-mediated posttranscriptional regulation. Furthermore, sX13 is the first sRNA shown to control the mRNA level of *hfq*, which encodes a conserved RNA-binding protein required for sRNA activity and virulence in many enteric bacteria.

the activity of the translational repressor protein RsmA, which impacts on QS, the production of extracellular enzymes and virulence [21,22,23]. In *Xcv*, sX12 was reported to be required for full virulence [16].

Xanthomonads are ubiquitous plant-pathogenic bacteria that infect approximately 120 monocotyledonous and 270 dicotyledonous plant species, many of which are economically important crops [24,25]. These pathogens, usually only found in association with plants and plant parts, differ from most other Gram-negative bacteria in their high G+C content (~65%), and high numbers of TonB-dependent transporters and signaling proteins [26]. Pathogenicity of most *Xanthomonas* spp. and other Gram-negative plant- and animal-pathogenic bacteria relies on a type III secretion (T3S) system which translocates bacterial effector proteins into the eukaryotic host cell [27,28]. In addition, other protein secretion systems, degradative enzymes and QS regulation contribute to virulence of *Xanthomonas* spp. [29,30].

One of the models to study plant-pathogen interactions is *Xcv*, the causal agent of bacterial spot disease on pepper and tomato [31,32]. The T3S system of *Xcv* is encoded by the *hrp−* [hypersensitive response (HR) and pathogenicity] gene cluster and translocates effector proteins into the plant cell where they interfere with host cellular processes to the benefit of the pathogen [29,33,34]. The mutation of *hrp*-genes abolishes bacterial growth in the plant tissue and the induction of the HR in resistant plants. The HR is a local and rapid programmed plant cell death at the infection site and coincides with the arrest of bacterial multiplication [33,35,36]. The expression of the T3S system is transcriptionally induced in the plant and in the synthetic medium XVM2, and is controlled by the key regulators HrpG and HrpX [37,38,39,40]. The OmpR-type regulator HrpG induces transcription of *hrpX* which encodes an AraC-type activator [39,41]. HrpG and HrpX control the expression of *hrp*, type III effector and other virulence genes [16,29,40,42,43]. Recently, dRNA-seq identified 24 sRNAs in *Xcv* strain 85-10, expression of eight of

which is controlled by HrpG and HrpX, including the aforementioned sX12 sRNA [15,16].

In this study, we aimed at a detailed characterization of sX13 from *Xcv* strain 85-10, which is an abundant and HrpG-/HrpX-independently expressed sRNA [16]. Using mutant and complementation studies, we discovered that sX13 promotes the expression of the T3S system and contributes to virulence of *Xcv*. Microarray and quantitative reverse transcription PCR (qRT-PCR) analyses identified a large sX13 regulon and G-rich motifs in presumed sX13-target mRNAs. Selected putative targets were analyzed by site-directed mutagenesis of *sX13* and mRNA::*gfp* fusions. Furthermore, we provide evidence that sX13 acts Hfq-independently. Our study provides the first comprehensive characterization of a *trans*-encoded sRNA that contributes to virulence of a plant-pathogenic bacterium.

Results

sX13 contributes to bacterial virulence

The sRNA sX13 (115 nt; [16]) is encoded in a 437-bp intergenic region of the *Xcv* 85-10 chromosome, i. e., 175 bp downstream of the stop codon of the DNA polymerase I-encoding gene *polA* and 148 bp upstream of the translation start site (TLS) of *XCV4199*, which encodes a hypothetical protein. Sequence comparisons revealed that the *sX13* gene is exclusively found in members of the *Xanthomonadaceae* family, i. e., in the genomes of plant-pathogenic *Xanthomonas* and *Xylella* species, the human pathogen *Stenotrophomonas maltophilia* and non-pathogenic bacteria of the genus *Pseudoxanthomonas*. Interestingly, *sX13* homologs are highly conserved [16] and always located downstream of *polA*. By contrast, *sX13*-flanking sequences are highly variable.

To characterize *sX13* in *Xcv* strain 85-10, we introduced an unmarked *sX13* deletion into the chromosome (see 'Materials and Methods'). Analysis of bacterial growth of the *sX13* mutant strain (Δ*sX13*) revealed a significantly reduced stationary-phase density compared to the *Xcv* wild-type strain 85-10 in complex medium (NYG; Figure 1A) and in minimal medium A (MMA; Figure 1B). The mutant phenotype of *Xcv*Δ*sX13* was complemented by chromosomal re-integration of *sX13* into the *sX13* locus, termed Δ*sX13*+*sX13*ₒₕ (Figure 1A, B; see 'Materials and Methods').

To address a potential role of *sX13* in virulence, we performed plant infection assays. As shown in Figure 1C, the *Xcv* strains 85-10 and Δ*sX13* grew similarly in leaves of susceptible pepper plants (ECW). Strikingly, infection with the *sX13* mutant resulted in strongly delayed disease symptoms in susceptible and a delayed HR in resistant pepper plants (ECW-10R) (Figure 1D). Ectopic expression of sX13 under control of the *lac* promoter (p*sX13*), which is constitutive in *Xcv* [38], and re-integration of *sX13* into the Δ*sX13* locus fully complemented the mutant phenotype of *Xcv*Δ*sX13* (Figure 1D; data not shown). Strain *Xcv* 85-10 carrying p*sX13* induced an accelerated HR in comparison to the wild type (data not shown).

Deletion of *sX13* derogates *hrp*-gene expression

As the HR induction in ECW-10R plants depends on the recognition of the bacterial type III effector protein AvrBs1 by the plant disease resistance gene *Bs1* [44,45], the *in planta* phenotype of *Xcv*Δ*sX13* suggested a reduced activity of the T3S system. To address this question, we investigated protein accumulation of selected components of the T3S system. Given that T3S apparatus proteins are not detectable in NYG-grown bacteria, we incubated the bacteria for 3.5 hours in the *hrp*-gene inducing XVM2 medium [38,40]. Western blot analysis revealed reduced amounts of the translocon protein HrpF, the T3S-ATPase HrcN and the
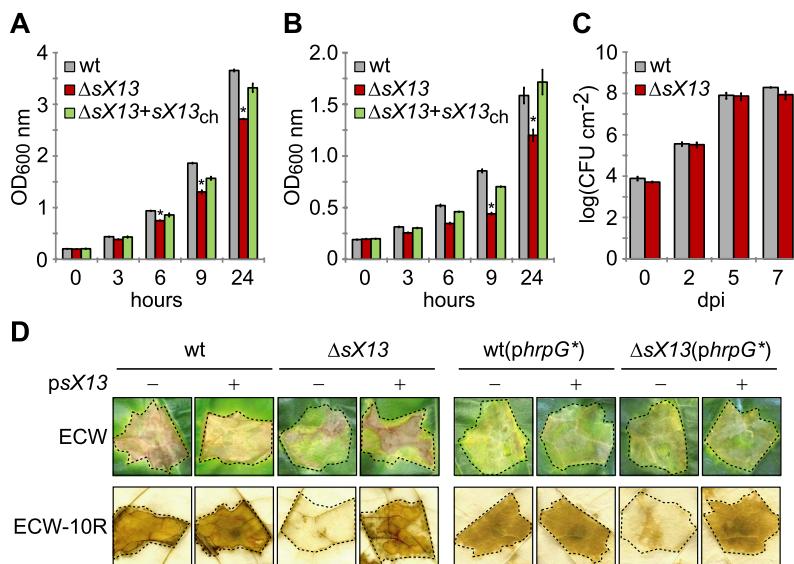
Figure 1. *sX13* contributes to bacterial growth in culture and virulence. Growth of *Xcv* wild type 85-10 (wt), the *sX13* deletion mutant (Δ*sX13*) and Δ*sX13* containing chromosomally re-integrated *sX13* (Δ*sX13*+*sX13*ch) in (A) complex medium NYG and (B) minimal medium MMA, respectively. Error bars represent standard deviations. Asterisks indicate statistically significant differences compared to wt (*t*-test; *P*<0.05). (C) Growth of *Xcv* 85-10 (wt) and Δ*sX13* in leaves of susceptible ECW pepper plants. Data points represent the mean of three different samples from three different plants of one experiment. Standard deviations are indicated by error bars. (D) Plant infection assay. *Xcv* strains 85-10 (wt) and Δ*sX13* carrying the empty vector (pB) or the sX13 expression construct (p*sX13*) and strains additionally expressing HrpG* (p*hrpG**) were inoculated at a density of 4×10^8 (left panel) and 10^8 cfu/ml (right panel), respectively, into leaves of susceptible ECW and resistant ECW-10R pepper plants. Disease symptoms in ECW were photographed 9 days post inoculation (dpi). The HR was visualized by ethanol bleaching of the leaves 3 dpi (left panel) and 18 hours post inoculation (right panel), respectively. Dashed lines indicate the inoculated areas. All experiments were performed at least three times with similar results.
doi:10.1371/journal.ppat.1003626.g001

T3S-apparatus component HrcJ in *Xcv*Δ*sX13* compared to the wild type, Δ*sX13*(p*sX13*) (Figure 2A) and strain Δ*sX13*+*sX13*ch (selectively tested for HrcJ; Figure 2B). Thus, sX13 positively affects the synthesis of T3S components.

As HrpG controls the expression of the *hrp*-regulon [39], we analyzed whether the reduced virulence of strain Δ*sX13* is due to a reduced activity of HrpG. Therefore, we ectopically expressed a constitutively active version of HrpG (HrpG*; p*hrpG**; [41]) in *Xcv*Δ*sX13* and performed plant-infection assays. The disease symptoms induced by *Xcv*Δ*sX13* and the wild type were comparable in presence of p*hrpG**, whereas with low inoculum of *Xcv* 85-10Δ*sX13* the HR was slightly delayed (Figure 1D). This suggests that HrpG* suppresses the 85-10Δ*sX13* phenotype. HrpF, HrcN and HrcJ protein accumulation in strain Δ*sX13*(p*hrpG**) was identical to the wild type suggesting full complementation (Figure 2A, B).

To investigate whether the reduced protein amounts of T3S-system components in *Xcv*Δ*sX13* are due to altered mRNA levels, we performed qRT-PCR analyses. mRNA accumulation of *hrpF*, *hrcJ* and the type III effector genes *avrBs1* and *xopJ* was two-fold lower in *Xcv*Δ*sX13* than in the wild type and the complemented strain Δ*sX13*+*sX13*ch (Figure 2C). In addition, the mRNA amount of *hrpX*, but not of *hrpG*, was reduced in the *sX13* mutant (Figure 2C). In presence of p*hrpG**, comparable mRNA amounts of *hrpG*, *hrpX*, *hrpF*, *hrcJ* and *xopJ* were detected in *Xcv* 85-10, Δ*sX13* and Δ*sX13*+*sX13*ch, whereas the *avrBs1* mRNA accumulation was significantly reduced in strain 85-10Δ*sX13* (Figure 2C). Taken together, our data suggest that the reduced virulence of the 85-10Δ*sX13* mutant is caused by a lower expression of components and substrates of the T3S system (Figure 1D; Figure 2A–C).

The deletion and chromosomal re-insertion of *sX13* in *Xcv*Δ*sX13* and Δ*sX13*+*sX13*ch, respectively, were verified by Northern blot using an sX13-specific probe (Figure S1). The sX13 abundance was not affected by expression of HrpG*, which confirms our previous findings [16] and suggests that expression of sX13 is independent of HrpG and HrpX (Figure S1).

sX13 accumulates under stress conditions

The expression of known bacterial sRNAs depends on a variety of environmental stimuli, which often reflect the physiological functions of sRNAs [2,46], e. g., the *E. coli* sRNA Spot42 is repressed in the absence of glucose and regulates carbon metabolism [47,48]. Northern blots revealed similar sX13 amounts in bacteria incubated in NYG medium at 30°C (standard condition), in presence of H_2O_2, at 4°C and in NYG medium lacking nitrogen (Figure 3A). By contrast, sX13 accumulation was increased in presence of high salt (NaCl), 37°C and in MMA (Figure 3A). Hence, sX13 is differentially expressed in different growth conditions and might contribute to environmental adaptation of *Xcv*.

Microarray analyses suggest a large sX13 regulon

To gain an insight into the sX13 regulon we performed microarray analyses. For this, cDNA derived from *Xcv* strains 85-10 and Δ*sX13* grown in NYG and MMA, respectively, was used as a probe. The hybridization data were evaluated using EMMA 2.8.2 [49] (see 'Materials and Methods'). In *Xcv*Δ*sX13* grown in NYG, 23 mRNAs were upregulated and 21 mRNAs were downregulated by a factor of at least 1.5 compared to the wild type (Table S2). In the MMA-grown *sX13* mutant, 23 upregulated
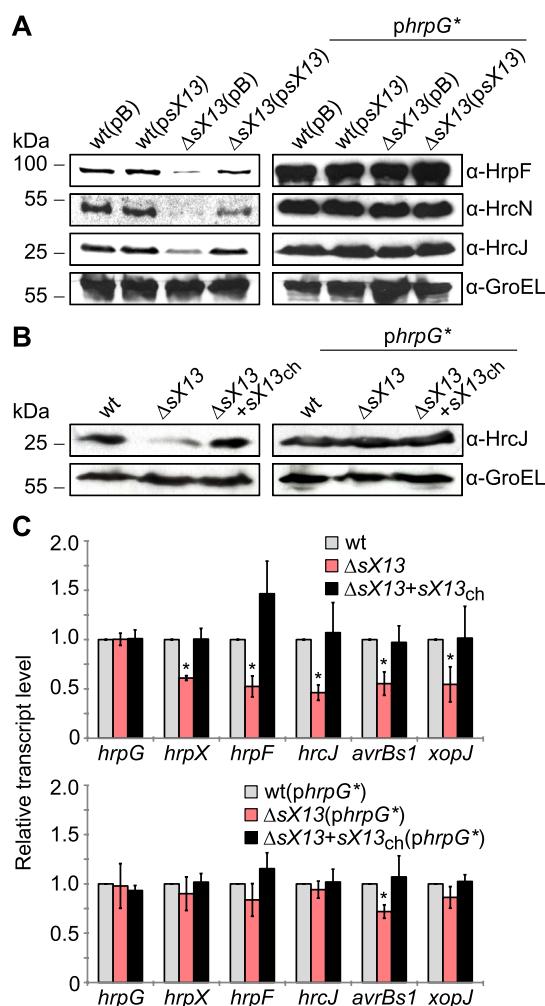
Figure 2. Deletion of *sX13* derogates virulence gene expression. (A) *Xcv* strains 85-10 (wt) and the *sX13* deletion mutant (Δ*sX13*) carrying the empty vector (pB) or the sX13 expression construct (p*sX13*) and strains additionally expressing HrpG* (p*hrpG**) were incubated for 3.5 hours in *hrp*-gene inducing medium XVM2. Total protein extracts were analyzed by immunoblotting using antibodies directed against HrpF, HrcN, HrcJ and GroEL. The experiment was repeated twice with similar results. (B) *Xcv* 85-10 (wt), Δ*sX13* and Δ*sX13*+*sX13*ch and strains additionally expressing HrpG* were incubated for 3.5 hours in *hrp*-gene inducing medium XVM2. Total protein extracts were analyzed by immunoblotting using antibodies directed against HrcJ and GroEL. The experiment was repeated twice with similar results. (C) Indicated genes were analyzed by qRT-PCR using RNA from cultures described in (B). The amount of each RNA in *Xcv* 85-10 was set to 1. Data points and error bars represent mean values and standard deviations obtained with three independent biological samples. Asterisks indicate statistically significant differences compared to wt (*t*-test; $P < 0.03$).
doi:10.1371/journal.ppat.1003626.g002

Figure 3. sX13 accumulation is altered under stress conditions in *Xcv* 85-10. (A) Northern blot analysis of sX13. Exponential phase cultures of NYG-grown *Xcv* 85-10 were transferred to NYG medium or MMA containing the indicated additives or lacking a nitrogen or carbon source (ΔN; ΔC). Cultures were shaken for three hours at 30°C unless otherwise indicated. 5S rRNA was probed as loading control. (B) sX13 and selected sX13-regulated genes (see Table 1) were analyzed by qRT-PCR using RNA from *Xcv* 85-10 (wt) cultures shown in (A) and NYG-grown Δ*sX13*. Bars represent fold-changes ($\log_{10}$) of mRNA amounts compared to *Xcv* 85-10 grown in NYG at 30°C. Experiments were performed twice with similar results.
doi:10.1371/journal.ppat.1003626.g003

and four downregulated genes confirmed the microarray data (Table 1; Figure 4).

sX13 negatively affects *hfq* and type IV pilus-biosynthesis mRNAs

Based on the annotated genome sequence of *Xcv* 85-10 [32], genes upregulated in *Xcv*Δ*sX13* can be grouped (Table S2): 18 genes encode proteins with unknown function, e. g., the putative LysM-domain protein XCV3927. 14 genes encode proteins involved in type IV pilus (Tfp) biogenesis, e. g., the putative Tfp assembly protein XCV2821, the pilus component PilE and the TCS response regulator PilG. Tfp enable twitching motility, i. e., adhesion to and movement on solid surfaces [50,51]. Three genes encode proteins assigned to signal transduction, i. e., the TCS regulator AlgR, the GGDEF-domain protein XCV2041 and the chemotaxis regulator XCV2186. Moreover, *hfq* mRNA accumulation was two-fold increased in *Xcv*Δ*sX13*.

The microarray data suggested that most upregulated genes in *Xcv*Δ*sX13* were only expressed in NYG- or MMA-grown bacteria (Table S2), which might be explained by the *P*-value and signal-strength thresholds applied for data evaluation. qRT-PCR analyses showed that the mRNA accumulation of *hfq*, *XCV2186*, *pilG* and *XCV3927* was increased in both the NYG- and MMA-

mRNAs were detected, four of which were also upregulated in NYG-grown bacteria, whereas no downregulated genes were identified (Table S2). With respect to both growth conditions, 42 and 21 genes were upregulated and downregulated, respectively, in the *sX13* mutant. qRT-PCR analyses of 11 selected upregulated
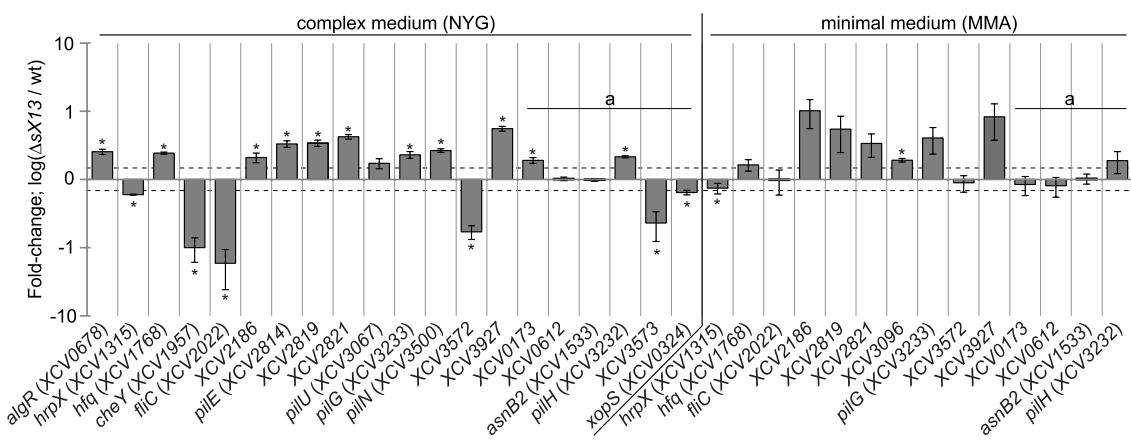
Figure 4. qRT-PCR analysis of sX13-regulated genes. Selected sX13-regulated genes (see Table 1) were analyzed by qRT-PCR using RNA from NYG- and MMA-grown *Xcv* strains 85-10 (wt) and Δ*sX13*. The amount of each mRNA in the wt was set to 1. Bars represent fold-changes of mRNA amounts in strain Δ*sX13* compared to 85-10 on a logarithmic scale ($\log_{10}$). Data points and error bars represent mean values and standard deviations obtained with at least three independent biological samples. Asterisks denote statistically significant differences (*t*-test; $P<0.05$). Dashed lines indicate a 1.5-fold change. Transcripts not detected in the microarray analyses are marked with 'a'.
doi:10.1371/journal.ppat.1003626.g004

grown *sX13* mutant compared to the wild type (Figure 4; Table 1). qRT-PCR analyses also revealed an upregulation of *pilH* in the NYG- and MMA-grown *Xcv*Δ*sX13* compared to the wild type (Figure 4; Table 1). Because *pilH* is the second gene in the *pilG* operon and was not detected as expressed in the microarray data, the number of mRNAs affected by *sX13* deletion might be higher than suggested by the microarray data.

sX13 positively affects *hrpX* and chemotaxis-regulating mRNAs

Five of 21 genes downregulated in *Xcv*Δ*sX13* presumably encode proteins involved in flagellum-mediated chemotaxis, e. g., the sensor kinase CheA1, the corresponding response regulator CheY and the flagellum components FliD and FliC (Table S2). qRT-PCR analyses revealed 17-fold lower *fliC* mRNA abundance in *Xcv*Δ*sX13* grown in NYG compared to the wild type, whereas the accumulation in MMA-grown cells was identical (Figure 4; Table 1). Similarly, *XCV3572*, which encodes a TonB-dependent receptor, was downregulated in NYG- but not in MMA-grown *Xcv*Δ*sX13* (Figure 4; Table 1). Gene *XCV3573*, which is encoded adjacent to *XCV3572* and encodes an AraC-type regulator, was also downregulated (Figure 4; Table 1). As mentioned above, sX13 positively affected the mRNA accumulation of *hrpX* in XVM2 medium (see Figure 2C), which was also true for bacteria grown in NYG and MMA (Figure 4; Table 1). Since HrpX controls the expression of many type III effector genes, we analyzed *xopS* [52] by qRT-PCR and detected similarly decreased levels in NYG-grown *Xcv*Δ*sX13* as for *hrpX* (Figure 4; Table 1). Taken together, our data suggest that the sX13 regulon comprises genes involved in signal transduction, motility, transcriptional and posttranscriptional regulation and virulence.

Accumulation of potential target mRNAs correlates with sX13 abundance

To address whether differential expression of sX13 under different conditions (see Figure 3A) affects the mRNA abundance of sX13-regulated genes, we performed qRT-PCR. We detected

elevated sX13 levels in *Xcv* strain 85-10 cultivated in high salt conditions, at 37°C and in MMA compared to standard conditions and an increased *hrpX* and decreased *XCV3927* mRNA accumulation (Figure 3B). In addition, low amounts of the *hfq* mRNA were detected in presence of high sX13 levels, whereas the abundance of the sX13-independent *XCV0612* mRNA (see Table 1) was not altered (Figure 3B).

sX13 activity does not require Hfq

The *hfq* mRNA accumulation in *Xcv*Δ*sX13* (Figure 3B; Figure 4; Table 1) prompted us to test whether sX13 activity depends on the RNA-binding protein Hfq. For this, we introduced a frameshift mutation into the *hfq* gene of *Xcv* strains 85-10 and 85-10Δ*sX13*. Northern blot analyses revealed comparable sX13 accumulation in both strains and the complemented *hfq* mutant, which ectopically expressed Hfq (p*hfq*) (Figure 5A). By contrast, the accumulation of the sRNA sX14 [16] was strongly reduced in the *hfq* mutant; this was restored by p*hfq* (Figure 5A). Unexpectedly, the *hfq* mutant strain was not altered in the induction of *in planta* phenotypes, i. e., in virulence (Figure 5B).

To investigate whether sX13 affects translation of putative target mRNAs, we established a GFP-based *in vivo* reporter system for *Xcv* similar to the one described for *E. coli* [53]. The promoterless broad-host range plasmid pFX-P permits generation of translational *gfp* fusions in a one-step restriction-ligation reaction (Golden Gate cloning [54]; see 'Materials and Methods'). We cloned the promoter, 5'-UTRs, and 10 and 25 codons of *XCV3927* and *hfq*, respectively, into pFX-P resulting in pFX*3927* and pFX*hfq*. *XCV3927* was selected because of a strongly increased mRNA accumulation in *Xcv*Δ*sX13* compared to the wild type (see Table 1). In presence of pFX*3927* or pFX*hfq*, fluorescence of XCV3927::GFP or Hfq::GFP fusion proteins was comparable in the *Xcv* wild type and *hfq* mutant (Figure 5C). The XCV3927::GFP and Hfq::GFP fluorescence was about 4-fold and 2-fold increased, respectively, in *Xcv*Δ*sX13* compared to strain 85-10 (Figure 5C), suggesting that the synthesis of the fusion proteins is repressed by sX13. Interestingly, the XCV3927::GFP and Hfq::GFP fluorescence was similarly increased in *Xcv*Δ*sX13* and the *sX13hfq* double

Table 1. Selected sX13-regulated genes validated by qRT-PCR analysis.

Locus[a]	Annotated gene product[b]	4G-motif[c]	Microarray – Fold-change (ΔsX13/wt)[d]		qRT-PCR – Fold-change (ΔsX13/wt)[e]	
			NYG	MMA	NYG	MMA
Upregulated genes (ΔsX13/wt)						
XCV0678	AlgR; two-component system regulatory protein	a,a,a	1.8	—	2.5±0.23	n.t.
XCV1768	Hfq; host factor-I protein	b	1.6	—	2.4±0.08	1.6±0.31
XCV2186	methyl-accepting chemotaxis protein	a	7.7	—	2.1±0.34	10.2±4.63
XCV2814	PilE; type IV pilus pilin	—	2.8	—	3.3±0.36	n.t.
XCV2819	type IV pilus assembly protein PilW	a	3.7	4.0	3.4±0.37	5.5±3.0
XCV2821	type IV pilus assembly protein FimT	a	4.3	7.4	4.2±0.32	3.4±1.27
XCV3067	PilU; type IV pilus assembly protein ATPase	a	1.8	—	1.7±0.29	n.t.
XCV3096	ComEA-related DNA uptake protein	—	—	4.2	n.t.	1.9±0.12
XCV3233	PilG; type IV pilus response regulator	a,b	—	2.0	2.3±0.26	4.1±1.71
XCV3500	PilN; type IV pilus assembly protein	—	2.7	—	2.7±0.16	n.t.
XCV3927	putative secreted protein	a	—	1.7	5.6±0.45	8.3±4.54
Downregulated genes (ΔsX13/wt)						
XCV1315	HrpX; AraC-type transcriptional regulator	—	0.6	—	0.6±0.01	0.7±0.13
XCV1957	CheY; chemotaxis response regulator	—	0.4	—	0.1±0.04	n.t.
XCV2022	FliC; flagellin and related hook-associated proteins	—	0.2	—	0.06±0.03	1.0±0.39
XCV3572	TonB-dependent outer membrane receptor	a	0.2	—	0.2±0.04	0.9±0.24
Additional genes tested by qRT-PCR						
XCV0173	putative secreted protein	a,b,b,b	—	—	1.9±0.19	0.8±0.26
XCV0612	ATPase of the AAA+ class	a	—	—	1.0±0.06	0.8±0.26
XCV1533	AsnB2; asparagine synthase	b	—	—	1.0±0.04	1.0±0.17
XCV3232	PilH; type IV pilus response regulator	a	—	—	2.2±0.07	1.9±0.67
XCV3573	putative transcriptional regulator, AraC family	a	—	—	0.2±0.11	n.t.
XCV0324	type III effector protein XopS	—	—	—	0.6±0.05	n.t.

[a], bold letters indicate genes with known TSS [16].
[b], refers to Thieme *et al.* (2005) [32].
[c], presence of a 4G-motif within the 5′-UTR or 100 bp upstream of translation start codon if TSS is unknown (a) and within 100 bp downstream of start codon (b) (see Figure S4).
[d], genes not detected as expressed are marked with —.
[e], values represent mean fold-change and standard deviation (see Figure 4);
n.t. - not tested.
doi:10.1371/journal.ppat.1003626.t001

mutant (Figure 5C). As abundance and activity of sX13 were not affected by the *hfq* mutation, we assume that sX13 acts Hfq-independently.

sX13 activity *in planta* depends on C-rich loop motifs

The predicted secondary structure of sX13 obtained by mfold [55] displays an unstructured 5′-region and three stable stem-loops, termed stem 1 to 3, and loop 1 to 3 (Figure 6A). Interestingly, loop 1 and loop 2 contain a 'CCCC' (4C) motif, whereas loop 3 harbors a 'CCCCC' (5C) motif (Figure 6A). To experimentally verify the predicted structure, we performed structure analyses of *in vitro* transcribed and radioactively-labeled sX13 treated with RNase V1 or RNase T1. While RNase T1 cleaves single-stranded RNA with a preference for G residues, RNase V1 randomly cleaves double-stranded RNA. We detected RNase T1-cleavage products for the 5′-region and RNase V1-cleavage products for stem 1 and 2, which is in good agreement with the predicted structure (Figure 6A; Figure S2). Moreover, RNase V1-cleavage products were less abundant for the 4C-motif of loop 1 and loop 2, suggesting single-stranded sequences

(Figure 6A; Figure S2). The results did not allow conclusions about stem 3 and loop 3 structures.

To assess the contribution of the 4C-/5C-motifs to sX13 function, we mutated p*sX13* in loop 1 and 2, respectively, to 'GCGC', and the 5C-motif in loop 3 to 'GCGCC' resulting in p*L1*, p*L2* and p*L3* (Figure 6A). In addition, loop mutations were combined (p*L1/2*, p*L1/3*, p*L2/3*) and analyzed for their ability to complement the *in planta* phenotype of strain Δ*sX13*. As shown above, *Xcv*Δ*sX13* induced a delayed HR, which was complemented by p*sX13* (Figure 1D). Similar phenotypes were observed with *sX13* mutants carrying p*L1* or p*sX13Δ5′*, which encodes a 5′-truncated sX13 derivative lacking the terminal 14 nucleotides (Figure 6B). The HR induced by the *sX13* mutant containing p*L2* or p*L1/2* was intermediate, whereas p*L3*, p*L1/3* and p*L2/3* failed to complement *Xcv*Δ*sX13* (Figure 6B). Northern blot analyses revealed expression of all *sX13*-loop mutant derivatives (Figure S3). The different RNA species derived from ectopically expressed sX13 and derivatives compared to chromosomally encoded sX13 might be due to alternative transcription termination of plasmid-derived sX13 and derivatives.

Figure 5. sX13 activity does not require Hfq. (A) Northern blot analysis. Total RNA from NYG-grown *Xcv* strains 85-10 (wt), the *hfq* frameshift mutant (*hfq*⁻) and the *hfq* mutant ectopically expressing Hfq (p*hfq*) was analyzed using sX13- or sX14-specific probes. 5S rRNA was probed as loading control. The experiment was performed twice with two independent mutants and with similar results. (B) Plant infection assay. The *Xcv* wild-type 85-10 (wt) and *hfq* mutant strain (*hfq*⁻) were inoculated at 2×10^8 cfu/ml into leaves of susceptible ECW and resistant ECW-10R plants. Disease symptoms were photographed 6 dpi. The HR was visualized 2 dpi by ethanol bleaching of the leaves. Dashed lines indicate the inoculated areas. The experiment was repeated three times with similar results. (C) GFP fluorescence of NYG-grown *Xcv* 85-10 (wt), the *hfq* mutant (*hfq*⁻), the *sX13* deletion mutant (Δ*sX13*) and the double mutant (Δ*sX13hfq*⁻) carrying pFX3927 or pFX*hfq*. *Xcv* autofluorescence was determined by *Xcv* 85-10 carrying pFX0 (control). Data points and error bars represent mean values and standard deviations obtained with at least four independent experiments. GFP fluorescence of the wt was set to 1. Asterisks denote statistically significant differences (*t*-test; $P < 0.01$).
doi:10.1371/journal.ppat.1003626.g005

sX13 loops differentially contribute to mRNA accumulation

As mutation of sX13 loops impinged on *Xcv* virulence (Figure 6B), we addressed by qRT-PCR whether loop mutations affect the mRNA abundance of *XCV2821*, *XCV3927*, *hfq* and *pilH*, which were upregulated in *Xcv*Δ*sX13* (see Figure 4; Table 1). In addition, we analyzed a downregulated gene, *XCV3572*, and *XCV0612*, which was not affected by *sX13* deletion. As shown in Figure 7A–E, sX13 negatively affected the mRNA abundance of *XCV2821*, *XCV3927*, *hfq* and *pilH*, whereas sX13 promoted mRNA accumulation of *XCV3572*. Mutation of *XCV3572*. Mutation of sX13 loops differentially affected the mRNA abundance of the tested genes: p*L2* and p*L1/2* failed to complement *Xcv*Δ*sX13* with respect to the mRNA abundance of *XCV2821*, *XCV3927* and *hfq* (Figure 7A–C). Intermediate mRNA amounts of *XCV3927* and *hfq* were detected in *Xcv*Δ*sX13* carrying p*L1/3* or p*L2/3* compared to pB and p*sX13* (Figure 7B, C). Taken together, the mRNA abundance of *XCV2821*, *XCV3927* and *hfq* appears to be mainly controlled by sX13-loop 2. In contrast, *pilH* mRNA accumulation appears to depend on multiple sX13 loops as only p*sX13* and p*L1* complemented *Xcv*Δ*sX13* (Figure 7D). The reduced mRNA amount of *XCV3572* in *Xcv*Δ*sX13* was complemented by p*L1* and p*L3* but not by p*L1/3* (Figure 7E), which suggests redundant

Figure 6. sX13 loops impact on *Xcv* virulence. (A) Secondary structure of sX13 based on prediction and probing (see Figure S2). sX13 consists of an unstructured 5′-, three double-stranded regions (S1; S2; S3) and three loops (loop 1–3). 4C-/5C-motifs are highlighted in red. Bold letters indicate unpaired bases and bars mark double-stranded regions deduced from structure probing. Mutations in loops are boxed, exchanged nucleotides are underlined. (B) Derivatives mutated in loops 2 and 3 fail to complement the plant phenotype of Δ*sX13*. Leaves of resistant ECW-10R plants were inoculated at 10^8 cfu/ml with *Xcv* 85-10 (wt) and Δ*sX13* carrying pBRS (pB), p*sX13* or one of the following derivatives: sX13 lacking 14 terminal nucleotides (p*sX13*Δ5′), sX13 mutated in single loops (p*L1*, p*L2*, p*L3*) or in two loops (p*L1/2*, p*L1/3*, p*L2/3*). The HR was visualized by ethanol bleaching of the leaves 2 dpi. Dashed lines indicate the inoculated areas. The experiment was performed four times with similar results.
doi:10.1371/journal.ppat.1003626.g006

roles of sX13-loops. In presence of p*L2*, p*L1/2* or p*L2/3* in *Xcv*Δ*sX13*, the *XCV3572* mRNA levels were intermediate compared to *Xcv*Δ*sX13* carrying pB or p*sX13* (Figure 7E). As expected, the mRNA abundance of *XCV0612* was identical in the different strains (Figure 7F).
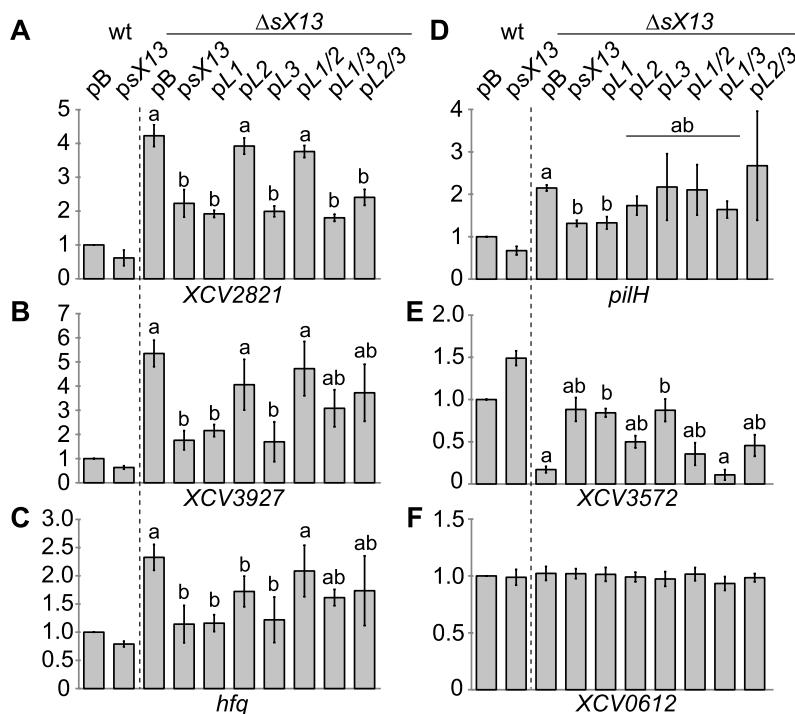
Figure 7. sX13 loops differentially contribute to abundance of putative mRNA targets. Relative transcript levels of (A) *XCV2821*, (B) *XCV3927*, (C) *hfq*, (D) *pilH*, (E) *XCV3572* and (F) *XCV0612* were analyzed by qRT-PCR in total RNA of NYG-grown *Xcv* strains 85-10 (wt) and Δ*sX13* carrying pBRS (pB), p*sX13* or mutated sX13-derivatives (see Figure 6). The mRNA abundance in the wt was set to 1. Data points and error bars represent mean values and standard deviations obtained with at least three independent biological samples. Statistically significant differences are indicated (*t*-test; $P<0.015$).
doi:10.1371/journal.ppat.1003626.g007

Identification of putative sX13-binding sites

To identify potential regulatory motifs in sX13-regulated mRNAs, a discriminative motif search was performed using DREME [56]. For this, sequences surrounding the TLSs of the 42 up- and 21 downregulated genes identified by microarray analyses (Table S2) were compared. More precisely, sequences spanning from known transcription start sites (TSSs) to 100 bp downstream of TLSs or, in case of unknown TSSs, 100 bp up- and 100 bp downstream of the TLS were inspected.

We found that up- and downregulated genes differ in the presence of 'GGGG' (4G) motifs. In the NYG-grown *sX13* mutant, 15 out of 23 (65%) upregulated genes contain up to three 4G-motifs which are predominantly located upstream of the TLS (Figure S4A; Table S2). 70% of the genes upregulated in MMA (16 out of 23), but only 14% of the genes downregulated in NYG medium (3 out of 21) contain 4G-motifs (Figure S4A; Table S2). Thus, 4G-motifs appear to be enriched in sX13-repressed mRNAs. However, the position of the motifs and flanking nucleotides are not conserved among sX13-regulated genes. Note that the term '4G-motif' also refers to motifs containing more than four G-residues in a row. The complementarity of C-rich sX13-loop sequences and G-rich mRNA motifs suggests sX13-mRNA interactions via antisense-base pairing (Figure 6A; Table 1; Table S2).

Compared to the occurrence of 4G-motifs in approximately 70% of sX13-repressed genes, only 30.71% of all chromosomally encoded *Xcv* genes (1,378 out of 4,487) carry 4G-motifs in

proximity of their TLS (Figure S4A). Interestingly, 4G-motifs in 241 of the chromosomally encoded genes (5.37%) are located between nucleotide position 8 and 15 upstream of the TLS (Figure S4B). This position corresponds to the presumed location of the RBS and suggests a role of 4G-motifs in translation control.

sX13 dependency of target::GFP synthesis requires both 4C- and 4G-motifs

To study the effect of sX13 on translation of selected putative targets, i. e., *XCV3927* and *hfq*, we used the above-mentioned GFP-reporter plasmids pFX*3927* and pFX*hfq*. In addition, we generated *pilH::gfp* (pFX*pilH*) and *XCV0612::gfp* (pFX*0612*) fusions (see 'Materials and Methods'). All mRNA::*gfp* fusions contain a G-rich motif in the proximity of their TLS which is complementary to C-rich sX13-loop regions (see 'Materials and Methods'). The fluorescence of the *sX13* deletion mutant carrying pFX*3927*, pFX*hfq* and pFX*pilH* was about 3.5-, 1.6- and 2.5-fold higher, respectively, compared to the *Xcv* wild type (Figure 8A–C). In presence of p*sX13*, pL1, pL3 or pL1/3 in *XcvΔsX13*, the XCV3927::GFP and Hfq::GFP fluorescence levels were comparable to the *Xcv* wild type (Figure 8A, B). By contrast, the XCV3927::GFP and Hfq::GFP fluorescence of strain Δ*sX13* carrying pL2, pL1/2 or pL2/3 was similarly increased as compared to *XcvΔsX13* carrying pB (Figure 8A, B). This suggests that the 4C-motif in sX13-loop 2 is required to repress XCV3927::GFP and Hfq::GFP synthesis. The increased PilH::GFP fluorescence of *XcvΔsX13* was complemented by
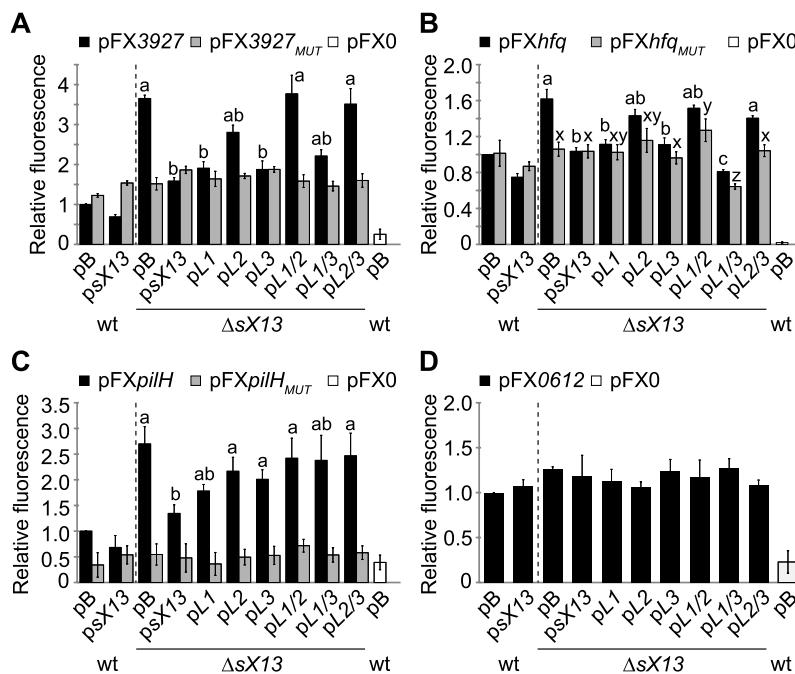
Figure 8. sX13-dependency of mRNA target::GFP synthesis requires a G-rich motif. GFP fluorescence of NYG-grown *Xcv* strains 85-10 (wt) and Δ*sX13* carrying pB, p*sX13* or mutated sX13-derivatives (see Figure 6) and carrying GFP-reporter plasmids (A) pFX*3927*, (B) pFX*hfq*, (C) pFX*pilH* or (D) pFX*0612*. pFX$_{MUT}$ derivatives contain a mutated 4G-motif. *Xcv* autofluorescence was determined using pFX0. GFP fluorescence of the wt was set to 1. Data points and error bars represent mean values and standard deviations obtained from at least four independent experiments. Statistically significant differences are indicated (*t*-test; *P*<0.015).
doi:10.1371/journal.ppat.1003626.g008

p*sX13* and p*L1*, in contrast to other *sX13*-loop mutant derivatives (Figure 8C). Fluorescence values of all analyzed *Xcv* strains carrying pFX*0612* were comparable confirming *sX13*-independency (Figure 8D).

To address whether the G-rich motif in presumed target mRNAs is required for sX13 dependency of mRNA::*gfp* translation, we introduced compensatory mutations. i. e., mutated the motif to 'GCGC'. *Xcv* strains carrying the resulting plasmids, pFX*3927$_{MUT}$*, pFX*hfq$_{MUT}$* or pFX*pilH$_{MUT}$*, exhibited a similar fluorescence in absence and presence of sX13 and mutated sX13 derivatives (Figure 8A–C). This suggests that the G-rich motif is required for sX13-dependency of target::GFP synthesis. However, mutation of the C-rich motifs in sX13 and the G-rich motifs in mRNA::*gfp* fusions did not restore sX13 dependency (Figure 8A–C). Unexpectedly, the fluorescence detected for *Xcv* strains containing pFX*3927$_{MUT}$* or pFX*hfq$_{MUT}$* was comparable to the fluorescence of *Xcv* 85-10 carrying the non-mutated plasmids pFX*3927* and pFX*hfq*, respectively (Figure 8A, B). The mutation of the 5G-motif in *pilH* abolished the fluorescence of strains containing pFX*pilH$_{MUT}$* indicating an essential role of the motif in *pilH* translation (Figure 8C).

Because sX13 was more abundant in MMA- than NYG-grown bacteria (Figure 3), we also analyzed the fluorescence of MMA-grown *Xcv* strains containing pFX-derivatives. XCV3927::GFP and PilH::GFP synthesis in MMA was sX13-dependently repressed to a greater extent than in NYG (Figure S5; see Figure 8).

Because sX13 negatively affected both the mRNA accumulation of chromosomally encoded *XCV3927*, *hfq* and *pilH* genes and accumulation of the corresponding GFP-fusion proteins, we exemplarily analyzed whether this is due to an altered mRNA abundance. However, qRT-PCR analyses revealed that the *XCV3927::gfp* mRNA accumulation was sX13-independent suggesting that sX13 posttranscriptionally affects the synthesis of XCV3927::GFP (Figure S6).

To discriminate between transcriptional and posttranscriptional effects of sX13 on target::GFP synthesis we generated reporter fusions controlled by p*lac* (see 'Materials and Methods'). Note that the activity of the *lac* promoter is not affected by deletion of *sX13* (data not shown). As shown in Figure S7, the fluorescence of *XcvΔsX13* carrying pFXpl-*3927* (*XCV3927*) and pFXpl-*pilH* (*pilH*) was 2.5- and 4-fold higher, respectively, compared to the *Xcv* wild type and the complemented *sX13* mutant strain. Interestingly, mutation of the 4G-motif in the *XCV3927* 5′-UTR did not only abolish sX13-dependency but also led to a significantly reduced fluorescence compared to the *Xcv* wild type which carried the non-mutated reporter plasmid (Figure S7). This suggests that the 4G-motif in the *XCV3927* 5′-UTR promotes translation, i. e., acts as translational enhancer element. In presence of pFXpl-*pilH*, the fluorescence of the fusion protein was only detectable in the *sX13* mutant but not in the wild type or complemented strain, confirming that PilH::GFP synthesis is repressed by sX13 (Figure S7). Overall, the data confirm that sX13 represses the synthesis of XCV3927 and PilH on the posttranscriptional level.

Discussion

sX13 controls *Xcv* virulence

This study provides a first insight into the posttranscriptional modulation of clade-specific adaptive processes in a plant-pathogenic γ-proteobacterium. We identified sX13 as a major regulator of *Xcv* virulence in that it promotes expression of genes in the *hrp*-regulon, i. e., components and substrates of the T3S system (Figure 2A–C). This finding is remarkable because the *hrp*-regulon is only expressed under certain conditions, whereas sX13 is constitutively expressed (Figure S1) [16]. The *sX13* gene is exclusively found and highly conserved in members of the *Xanthomonadaceae* family of Gram-negative bacteria [16]. Intriguingly, several species with an *sX13* homolog lack a T3S system, e. g., the plant pathogen *X. fastidiosa* and the opportunistic human pathogen *S. maltophilia*. This suggests a role of sX13 apart from regulation of the *hrp*-regulon in these organisms.

The expression of the *hrp*-regulon depends on HrpG and HrpX [39,40]. HrpG is presumably posttranslationally activated in the plant and in XVM2 medium and induces the expression of *hrpX* [38,39,40,41]. As the XVM2-grown *sX13* mutant displayed decreased mRNA amounts of *hrpX* but not of *hrpG* (Figure 2C), we suppose that sX13 acts upstream of HrpG. This idea is supported by the finding that constitutively active HrpG (HrpG* [41]) suppressed the *sX13* mutation with respect to virulence and the expression of *hrpX* and downstream genes (Figure 1D; Figure 2A–C). In addition, sX13 affected the basal expression level and, hence, the activity of HrpX under non-inducing conditions, which might impact on the efficiency of *hrp*-gene induction during infection. Based on the fact that HrpG::GFP and HrpX::GFP synthesis was sX13-independent (Figure S8) we assume that sX13 indirectly controls the expression of the *hrp*-regulon.

Physiological roles of sX13

Deletion of *sX13* affected the mRNA abundance of more than 60 genes involved in signaling, motility, transcriptional and posttranscriptional regulation (Table S2). sX13 negatively regulated mRNAs involved in Tfp biogenesis but promoted the accumulation of mRNAs involved in flagellum-mediated chemotaxis in a growth-phase dependent manner (Table 1; Table S2). This, together with the fact that sX13 is differentially expressed under certain stress conditions (Figure 3), implies a central role of sX13 in the transduction of environmental signals into comprehensive cellular responses affecting virulence gene expression, motility and QS regulation. The latter is corroborated by the reduced stationary-phase cell density of the *sX13* mutant compared to the *Xcv* wild type (Figure 1A, B) and the sX13-dependency of the *XCV2041* mRNA (Table S2), which encodes a GGDEF-/EAL-domain protein. Such domains play a role in the control of cyclic-di-GMP levels and QS regulation [57]. Interestingly, XCV2041 shares 94% identity with the *Xcc* protein XC2226 which is a repressor of Tfp-mediated motility [58].

Another remarkable finding of our study was the sX13-dependent accumulation of the *hfq* mRNA. To the best of our knowledge, sX13 is the first sRNA which affects expression of this conserved RNA-binding protein (Table 1). The *Xcv hfq* mutant was unaltered in virulence on its host plant (Figure 5B), which is in good agreement with recent findings for *Xoo* [18]. By contrast, Hfq contributes to virulence in a number of other bacteria, including the plant pathogen *A. tumefaciens*, and is also involved in symbiotic plant interactions of *Sinorhizobium meliloti* [5,12,59,60]. In *Vibrio cholerae*, four redundantly acting and Hfq-dependent sRNAs (Qrr) destabilize *hapR* mRNA, which encodes the master regulator of

QS, the T3S system and other virulence genes [61,62]. In the Gram-positive human pathogen *Staphylococcus aureus*, the Hfq-independent RNAIII is induced by the *agr* QS system and mediates the switch between the expression of surface proteins and excreted toxins through translational repression of Rot (repressor of toxins) [63,64,65].

sX13 activity depends on C-rich loop regions

Xcv sRNAs are strongly structured and lack extended single-stranded regions [15,16,32]. In contrast, enterobacterial sRNAs commonly exhibit a modular structure consisting of a single-stranded mRNA-targeting domain, often located at the 5′-end, an A/U-rich Hfq-binding site and a Rho-independent terminator [1]. The sX13 structure suggests that direct sRNA-mRNA interactions are energetically confined to the unstructured 5′-region and its three C-rich loops (Figure 6A). However, the 5′-region of sX13 was dispensable for full virulence of *Xcv* and sX13 activity appears to be exerted via loops 2 and 3 (Figure 6B). Although loops 1 and 2 just differ in the 3′-adjacent nucleotide (U/A) (Figure 6A), only loop 2 was required to repress the synthesis of XCV3927::GFP and Hfq::GFP, which might depend on the position of stem-loops in the sRNA and, thus, accessibility. By contrast, repression of PilH::GFP appears to depend on multiple sX13 regions (Figure 8C).

An important question is whether sX13 controls target gene expression on the level of mRNA stability or translation. On one hand, sX13-loop mutant derivatives affected the mRNA levels of presumed targets (Figure 7). On the other hand, protein levels, but not the mRNA level of an *XCV3927::gfp* fusion, harboring only the 5′-UTR and 10 codons of *XCV3927*, were sX13-dependent (Figure 8A; Figure S6). This suggests that the impact of sX13 on *XCV3927* mRNA abundance and translation are separate events and hints at the presence of additional regulatory sites in the *XCV3927* mRNA. It should be noted that the assessment of RNA stability by rifampicin treatment is hampered by the fact that our *Xcv* strains are rifampicin resistant.

The sX13 loops are reminiscent of regulatory RNAs in *S. aureus*, many of which contain 'UCCC'-motifs in loops [66]. For example, RNAIII contains C-rich stem-loops, which interact with the RBS of target mRNAs [63,65,67]. RNAIII represses Rot synthesis through formation of kissing complexes between two stem-loops of each RNAIII and *rot* mRNA [64,65]. Such multiple loop interactions are also employed by the *E. coli* sRNA OxyS to target *fhlA* [68]. In *Helicobacter pylori*, the sRNA HPnc5490 represses the synthesis of the chemotaxis regulator TlpB [69]. Interestingly, the central part of the HPnc5490-loop sequence is identical to the 'UCCCCCU'-motif of loop 3 in sX13 [69].

G-rich enhancer motifs confer sX13-dependency of target mRNAs

Similarly to RNAIII targets in *S. aureus* and the *tlpB* mRNA in *H. pylori* [65,69], mRNAs repressed by sX13 are enriched for G-rich motifs in proximity of the TLS (Figure S4; Table S2). The complementarity between these motifs and the 4C-/5C-motif in the sX13 loops suggests sX13-mRNA interactions through antisense base pairing. Our data emphasize that sX13 acts posttranscriptionally on target genes that contain G-rich motifs, as shown for *XCV3927* and *pilH* (Figure S7). This idea is supported by the fact that mutation of the G-rich motifs, located in the mRNA of *XCV3927*, *pilH* and *hfq*, abolished sX13-dependency of protein synthesis (Figure 8; Figure S5; Figure S7). However, the presence of a G-rich motif does not necessarily confer regulation by sX13 (see *XCV0612*; Figure 7F; Figure 8D). Given that eight of 28 repressed and 4G-motif-containing mRNAs contain at least

two 4G-motifs close to the TLSs (Figure S4), we assume that sX13 loops can interact with multiple 4G-motifs in certain mRNAs. As positively regulated mRNAs lack G-rich motifs, sX13 presumably acts indirectly on the corresponding genes (Figure S4; Table S2).

Direct sRNA-mRNA interactions are commonly validated by compensatory mutant studies [1]. However, in case of the *E. coli* sRNA RyhB, mutations were suggested to interfere with Hfq-binding and rendered compensatory mutants non-functional [70]. Here, mutation and deletion of *sX13* increased the synthesis of XCV3927::GFP and Hfq::GFP fusions, whereas mutation of corresponding 4G-motifs resulted in similar fluorescence values as non-mutated mRNA::*gfp* fusions in *Xcv* wild type. In addition, the reduced fluorescence of mutated target::GFP fusions was unaffected by compensatory sX13-mutant derivatives (Figure 8). This suggests that G-rich motifs in sX13-repressed mRNAs play a role besides mediation of sRNA interactions. While *Xanthomonas* spp., like other G+C-rich bacteria, lack a consensus RBS [16,71], 5% of the chromosomal *Xcv* coding sequences (241 of 4,487) contain a G-rich motif 8–15 nucleotides upstream of their TLS (Figure S4). As anticipated, mutation of the 5G-motif at the RBS position of *pilH* abolished translation (Figure 8; Figure S5; Figure S7). By contrast, the 4G-motifs in *XCV3927* and *hfq* mRNAs, located 21 nucleotides upstream and nine nucleotides downstream of the AUG, respectively, confer sX13-dependency but were not essential for translation (Figure 8). Thus, G-rich motifs confer sX13-dependency and mRNA translation in a position-dependent manner. As mutation of the 4G-motif in *XCV3927* reduced protein synthesis, the motif appears to function as translational enhancer (Figure S7). We suggest that sequestration of a G-rich motif by sX13 as well as mutation of the motif precludes the binding of an unknown factor, which promotes mRNA translation. Such a factor could be RNA, protein or the ribosome.

The presumed sX13 mode of action is reminiscent of the *Salmonella* sRNA GcvB, which inhibits translation of mRNAs by targeting C/A-rich enhancer elements [72,73]. By increasing the ribosome-binding affinity, C/A-rich motifs enhance mRNA translation, irrespective of their localization upstream or downstream of the TLS [72,74].

sRNAs encoded at the *polA* locus in other bacteria

Homologs of *Xcv* sRNAs are predominantly found in members of the *Xanthomonadaceae* family but not in other bacteria [15,16]. The *sX13* gene is located adjacent to the DNA polymerase I-encoding *polA* gene, which resembles a locus encoding the Spot42 sRNA in *E. coli* and members of the αr7 sRNA family in α-proteobacteria [75,76,77,78]. In contrast to sX13, Spot42 requires Hfq and regulates targets involved in carbon metabolism [48,79], e. g., the discoordinated expression of genes within the *gal* galactose utilization operon [47], which is absent in *Xcv*. Although sX13 lacks sequence similarity to Spot42 or αr7 sRNAs, the latter also contain three stem-loops and apical C-rich motifs [80] suggesting that sRNAs in distantly related bacteria evolved divergently but retained structural conservation. Thus, it will be interesting to see whether the *polA* locus in other bacteria also encodes sRNAs, and whether sX13 and structurally related sRNAs act in a similar manner.

Materials and Methods

Bacterial strains and growth conditions

For bacterial strains, plasmids and oligonucleotides used in this study see Table S1. *E. coli* strains were grown at 37°C in lysogeny broth (LB), *Xcv* strains at 30°C in nutrient-yeast-glycerol (NYG) [81], XVM2 [40] or minimal medium A (MMA) [82], which was supplemented with casamino acids (0.3%) and sucrose (10 mM). Plasmids were introduced into *E. coli* by chemical transformation and into *Xcv* by tri-parental conjugation, using pRK2013 as helper plasmid [83]. Antibiotics were added to the final concentrations: ampicillin, 100 μg/ml; gentamycin, 15 μg/ml; kanamycin, 25 μg/ml; rifampicin, 100 μg/ml; spectinomycin, 100 μg/ml.

Generation of constructs

To generate the sRNA-expression vector pBRS, a 28-bp fragment between the *lac* promoter and the *Eco*RI cloning site in pBBR1mod1 [84] was replaced by a truncated fragment, amplified by PCR from pBBR1mod1 using primers pBRS-EcoRI-fw and pBRS-NcoI-rev. For generation of constructs expressing sX13 (p*sX13*) and sX13Δ5′ (p*sX13Δ5′*; lacking 14 nt at the 5′-end), respective fragments were PCR-amplified from *Xcv* 85-10 using primers sX13-fw/-rev or sX13d5-fw/-rev. PCR products were digested with *Eco*RI/*Hin*dIII and cloned into pBRS. The *sX13*-mutant plasmids p*L1*, p*L2*, p*L3* and p*L2/3* were generated by PCR amplification of plasmid p*sX13* using primers L1-fw/-rev, L2-fw/-rev, L3-fw/-rev and L3-fw/L2/3-rev, respectively; plasmid p*L1/3* was generated with primers L3-fw/-rev and p*L1* as template; p*L1/2* was generated with primers L2-fw/L1/2-rev and p*L2* as template. For ectopic expression of *hfq* under control of its own promoter, a PCR product obtained from *Xcv* 85-10 using primers pMphfq-fw/-rev was cloned into the promoterless vector pBRM-P [84].

The GFP-based promoter-less reporter plasmid pFX-P permits *Bsa*I-mediated cloning of PCR amplicons (Golden Gate cloning) in a one-step restriction-ligation reaction [54] and was generated as follows: pDSK602 [85] was digested with *Pst*I/*Bam*HI to remove the *lac* promoter and multiple-cloning site. To allow blue-white selection, a dummy module containing 5′- and 3′-*Bsa*I recognition sites, p*lac* and *lacZ* was PCR-amplified from pBRM-P [84] using primers pFX-lz-fw/-rev. A fragment containing both the *gfp* coding sequence without translation start codon and the *rrnB* terminator was PCR-amplified from pXG-1 [53] using primers pFXgfp-fw/-rev. After blunt-end ligation of dummy- and *gfp*-module, the fragment was digested with *Mph1103*I/*Bgl*II and ligated into the *Pst*I/*Bam*HI sites of the pDSK602 backbone, resulting in pFX-P. For generation of the GFP-control plasmid pFX0, a promoterless fragment (138 bp) of the sRNA gene *sX6* [16] was PCR-amplified from *Xcv* 85-10 using primers pFX0-fw/-rev and cloned into pFX-P.

To generate mRNA::*gfp* expression constructs, fragments containing the promoter, 5′-UTR and 10 to 25 codons of the respective genes were PCR-amplified from *Xcv* 85-10 using corresponding pFX-fw/-rev primers (Table S1) and cloned into pFX-P. Plasmids pFX*3927*, pFX*hfq* and pFX*0612* were generated by cloning of nucleotide sequences −98 to +30, −160 to +75 and −116 to +33 relative to the translation start codon of *XCV3927*, *hfq* and *XCV0612*, respectively. pFX*pilH* was constructed by cloning a fragment spanning nucleotides −99 upstream of the *pilG* translation start codon to nucleotide +60 within the coding sequence of *pilH*.

The 4G-motif in *XCV3927::gfp* and *hfq::gfp* is located 21–24 bp upstream and 9–12 bp downstream of the ATG, respectively. *pilH::gfp* and *XCV0612::gfp* contain a 5G-motif at nucleotide positions 10–14 and 8–12 upstream of the ATG, respectively. Plasmids pFX$_{MUT}$ were constructed as follows: to mutate the 'GGGG' motif to 'GCGC', sequences upstream and downstream of the motif were PCR-amplified from *Xcv* 85-10 using primers pFX-fw/pFX-mut-L-rev and pFX-mut-R-fw/pFX-rev, respectively. Primers pFX-mut-L-rev and pFX-mut-R-fw contain the mutation flanked by a *Bsa*I-recognition site. pFX and correspond-

ing pFX$_{MUT}$ derivatives only differ in the sequence of the G-rich motif at nucleotide positions relative to the translation start codons: $-24/-22$ in pFX3927_{MUT}, $+10/+12$ in pFXhfq_{MUT} and $-12/-10$ in pFX$pilH_{MUT}$.

Plasmids pFXpl, which express p*lac*-driven mRNA::*gfp* fusions, were constructed by cloning respective fragments into pFX-P: p*lac* was PCR-amplified from pBRM-P [84] using primers plac-fw/rev; sequences -59 to $+54$ and -147 to $+54$ relative to the ATG of *hrpX* and *hrpG*, respectively, were PCR-amplified from *Xcv* 85-10 using primers pFXpl-hrpX-fw/-rev and pFXpl-hrpG-fw/-rev; fragments of *XCV3927* and *pilH* were PCR-amplified from respective pFX and pFX$_{MUT}$ plasmids using primers pFXpl3927-fw/pFXpl3927mut-fw/pFX3927-rev and pFXplpilH-fw/pFXpilH-rev.

Generation of *Xcv* mutant strains

To generate a chromosomal *sX13* deletion mutant, flanking sequences of ~650 bp up- and downstream of *sX13* [16] were amplified by PCR from *Xcv* 85-10 using primers d13L-fw/-rev and d13R-fw/-rev. PCR products were digested with *Bam*HI/*Hin*dIII and *Hin*dIII/*Xba*I, respectively, and ligated into the suicide vector pOK1 [86]. *Xcv*Δ*sX13*+*sX13*$_{ch}$, which carries an *sX13* copy at the Δ*sX13* locus, was created as follows: two PCR fragments amplified from *Xcv* 85-10 using primers int13L-fw/rev and int13R-fw/rev were digested with *Psp*1406I, ligated and cloned into the *Bam*HI/*Xba*I site of pOK1. Conjugation of pOKΔ*sX13* into *Xcv* 85-10 and pOKint13 into *Xcv*Δ*sX13* and selection of the correct double crossing-over events were performed as described [86]. *Xcv*Δ*sX13*+*sX13*$_{ch}$ was identified by PCR amplification of the *sX13* locus and *Psp*1406I restriction.

To introduce a frameshift mutation into chromosomal *hfq*, PCR products obtained from *Xcv* 85-10 using primers hfqL-fw/-rev and hfqR-fw/-rev were digested with *Bam*HI/*Bsa*I and *Bsa*I/*Xba*I, respectively, and cloned into pOK1. Conjugation of pOK-fs*hfq* into *Xcv* and selection of double crossing-over events were performed as described [86]. The resulting *hfq* mutant strain carries a 4 bp deletion in an *Mnl*I site (nucleotides 33–36 in *hfq*) and was identified by PCR using primers seqhfq-fw/-rev followed by digestion with *Mnl*I.

Plant material and plant inoculations

Pepper (*Capsicum annuum*) plants of the near-isogenic cultivars ECW and ECW-10R [87] were grown at 25°C with 60–70% relative humidity and 16 hours light. For infection assays, *Xcv* bacteria were resuspended in 10 mM MgCl$_2$ and inoculated with a needleless syringe into the intercellular spaces of leaves using concentrations of 1–4×10^8 colony-forming units (CFU) per ml for scoring plant reactions and 10^4 CFU/ml for *in planta* growth curves. For better visualization of the HR, leaves were bleached in 70% ethanol. *In planta* growth curves were performed as described [33]. All experiments were repeated at least two times.

Protein detection and measurement of GFP fluorescence in *Xcv*

Xcv cells grown overnight in NYG medium were washed, inoculated at OD$_{600}$ = 0.2 into XVM2 medium and incubated for 3.5 hours at 30°C. Total cell extracts were analyzed by sodium dodecyl sulfate-polyacrylamide gel electrophoresis and immuno-blotting using specific polyclonal antibodies directed against HrpF [88], HrcN [89], HrcJ [89] and GroEL (Stressgen). A horseradish peroxidase-labeled anti-rabbit antibody (Amersham Pharmacia Biotech) was used as secondary antibody. Antibody reactions were visualized by enhanced chemiluminescence (Amersham Pharmacia Biotech).

To determine GFP fluorescence, bacteria were adjusted to OD$_{600}$ = 1.0 in 10 mM MgCl$_2$. Fluorescence was measured at 485-nm excitation and 535-nm emission using a microplate reader (SpectraFluor Plus; Tecan Trading AG).

In vitro transcription and structure probing

sX13 [16] was PCR-amplified from *Xcv* 85-10 using primers sX13T7-fw, containing the T7-promoter, and sX13T7-rev and cloned into pUC57 (Thermo Fisher Scientific), resulting in pUC-13T7. Template DNA for *in vitro* transcription was amplified from pUC-13T7 using primers sX13-ITC-fw/-rev. *sX13* transcription and DNase treatment were performed according to manufacturer's instructions (MEGAscript®Kit; Invitrogen). RNA labeling using [γ-^{32}P]-ATP, treatment with RNase T1 (1 Pharmacia unit; Ambion) or RNase V1 (0.01 to 0.0002 Pharmacia units; Ambion) and generation of nucleotide ladders were performed as described [90]. Samples were analyzed on 12% polyacrylamide gels containing 7 M urea. Signals were visualized with a phosphoimager (FLA-3000 Series; Fuji).

RNA preparation, Northern blot and qRT-PCR analysis

Bacteria were grown overnight in NYG and inoculated at OD$_{600}$ = 0.2 into NYG or XVM2 medium. XVM2 cultures were incubated for 3.5 hours at 30°C. NYG-grown cells were harvested at exponential growth phase (OD$_{600}$ = 0.5–0.7) or used to inoculate the following media at OD$_{600}$ = 0.5: NYG containing 0.3 M NaCl, 0.2 M H$_2$O$_2$ or NYG lacking a nitrogen source, MMA or MMA lacking a carbon source followed by incubation for 3 hours.

RNA isolation and Northern blot hybridization was performed as described [16,91]. Oligonucleotide probes for detection of sX13 and 5S rRNA are described in [16].

For qRT-PCR analyses, cDNA was synthesized using Rever-tAid H Minus First Strand cDNA-Synthesis Kit according to manufacturer's instructions (Fermentas). qRT-PCR was performed using 2 ng cDNA and ABsolute BlueSYBR Green Fluorescein (Thermo Scientific) and analyzed on MyiQ2 and CFX Connect systems (Bio-Rad). The efficiency and specificity of PCR amplifications was determined by standard curves derived from a dilution series of template cDNA and melting curve analysis, respectively. Mean transcript levels were calculated based on values obtained from technical duplicates of at least three independent biological replicates and the levels of 16S rRNA (reference gene) as described (ABI user bulletin 2; Applied Biosystems).

Microarray analysis

For isolation of total RNA, NYG-grown cells were harvested at exponential growth phase (OD$_{600}$ = 0.5–0.7) or used to inoculate MMA at OD$_{600}$ = 0.5 followed by incubation for 3 hours. Fluorescently labeled cDNA was prepared as described [92]. Starting from 10–15 μg total RNA, aminoallyl-modified first strand cDNA was synthesized by reverse transcription using random hexamer primers, Bioscript reverse transcriptase (Bioline) and 0.5 mM dNTP, dTTP:aminoallyl-dUTP (1:4). After hydro-lysis and clean up using Nucleotide removal kit (Qiagen), Cy3- and Cy5-N-Hydroxysuccinimidyl ester dyes (GE Healthcare) were coupled to the aminoallyl-labeled first strand cDNA. Uncoupled dye was removed using the Nucleotide removal kit (Qiagen). For RNA from NYG- and MMA-grown bacteria, four and three microarray hybridizations were performed, respectively, using labeled cDNA obtained from independent bacterial cultures.

The genome-wide microarray for *Xcv* strain 85-10 (Xcv5KOLI) carried 50–70 nt unique oligonucleotides representing CDSs, with

each oligonucleotide spotted in three technical replicates per microarray [93]. Preprocessing of microarrays was performed as described [94]. Hybridization was performed in EasyHyb hybridization solution (Roche) supplemented with sonicated salmon sperm DNA at 50 μg/ml in a final volume of 130 μl for 90 min at 45°C using the HS400 Pro hybridization station (Tecan Trading AG). Labeled cDNA samples were denatured for 5 min at 65°C prior hybridization. After hybridization microarrays were washed as described [94].

Mean signal and mean local background intensities were obtained for each spot on the microarray images using ImaGene 8.0 software for spot detection, image segmentation and signal quantification (Biodiscovery Inc.). Spots were flagged as empty if $R \leq 0.5$ in both channels, where $R = (\text{signal mean} - \text{background mean})/\text{background standard deviation}$. Remaining spots were analyzed further. The $\log_2$ value of the ratio of intensities was calculated for each spot using the formula $M_i = \log_2(R_i/G_i)$. $R_i = I_{ch1(i)} - Bg_{ch1(i)}$ and $G_i = I_{ch2(i)} - Bg_{ch2(i)}$, where $I_{ch1(i)}$ or $I_{ch2(i)}$ is the intensity of a spot in channel 1 or channel 2, and $Bg_{ch1(i)}$ or $Bg_{ch2(i)}$ is the background intensity of a spot in channel 1 or channel 2. The mean intensity was calculated for each spot, $A_i = \log_2(R_iG_i)^{0.5}$ [95]. For data normalization (Median), significance test (Holm) and t-statistics analysis, the EMMA 2.8.2 open source platform was used [49]. Genes were accounted as differentially expressed if $P\ adjusted \leq 0.05$, $A \geq 8$, and if the ratio showed at least a 1.5-fold difference between the two experimental conditions.

Biocomputational analyses

Homology searches were performed using BLASTn and the NCBI genome database (http://blast.ncbi.nlm.nih.gov; http://www.ncbi.nlm.nih.gov/genome; date: 11/22/2012).

The secondary structure of sX13 [16] was predicted using Mfold version 3.5 (http://mfold.rna.albany.edu/?q = mfold/RNA-Folding-Form) and default folding parameters [55]. To identify putative regulatory motifs in the 5′-regions of sX13-regulated mRNAs, a discriminative motif search was performed using DREME version 4.9.0 (http://meme.nbcr.net/meme/cgi-bin/dreme.cgi) [56]. Sequences of regulated genes comprising nucleotide positions −100 to +100 relative to translation start codons or in case of known TSSs [16] (see Table S2), sequences comprising the 5′-UTR to position +100 downstream of translation start codons, were extracted from the genome of *Xcv* strain 85-10 (NC_007508 and NC_007507) [32]. DREME motif search was performed with negatively regulated genes as input and positively regulated genes as comparative sequences and an *E*-value of ≤5.

Accession numbers

YP_363045.1; YP_363046.1; YP_362142.1; YP_362163.1; YP_362160.1; YP_361663.1; YP_363499.1; YP_365931.1; YP_363887.1; YP_365930.1; YP_365658.1; YP_364552.1; YP_363772.1; YP_363917.1; YP_364964.1; YP_364963.1; YP_364545.1; YP_365303.1; YP_365304.1; YP_362343.1; YP_362302.1; YP_362409.1; YP_364550.1; YP_364798.1; YP_364827.1; YP_365231.1; YP_363688.1; YP_363753.1; YP_361904.1; YP_363264.1; YP_362055.1; YP_363957.1

Supporting Information

Figure S1 sX13 abundance is not affected by expression of HrpG*. *Xcv* 85-10 (wt), Δ*sX13* and Δ*sX13*+*sX13*ch and strains additionally expressing HrpG* were incubated for 3.5 hours in *hrp*-gene inducing medium XVM2 (see Figure 2B). Total RNA was analyzed by Northern blot using an sX13-specific probe. 5S rRNA

was probed as loading control. The experiment was performed twice with similar results.
(EPS)

Figure S2 Structure probing of sX13. *In vitro* transcribed sX13 was 5′-labeled and treated with RNase T1 (T1) or alkaline hydroxyl (OH⁻) buffer for generation of nucleotide ladders and RNase V1 (V1) for mapping of base-paired regions. Lane 'C' contains untreated sX13; triangle indicates decreasing concentrations of RNase V1; '#G' indicates positions of G residues; the deduced secondary structure is indicated on the right hand side (see Figure 6A).
(EPS)

Figure S3 Expression of sX13 derivatives. Total RNA of NYG-grown *Xcv* strains 85-10 (wt) and Δ*sX13* carrying pBRS (pB), p*sX13* or expressing mutated sX13-derivatives (see Figure 6) was analyzed by Northern blot using an sX13-specific probe. 5S rRNA was probed as loading control. The experiment was performed twice with similar results.
(EPS)

Figure S4 Distribution of 4G-motifs among sX13-regulated genes and chromosomally encoded *Xcv* genes. (A) Percentage of sX13-regulated genes identified by microarray analyses (see Table S2) and chromosomal CDSs in *Xcv* containing one or more 4G-motifs in region −100 to +100 relative to the TLS or in case of known TSSs, in the sequence comprising the 5′-UTR to position +100. The number of genes analyzed (n) is given below. (B) Distribution of 4G-motifs found in region −100 to +100 bp relative to the TLSs of 1,378 chromosomal CDSs [see (A)].
(EPS)

Figure S5 sX13-dependency of mRNA target::GFP synthesis in MMA-grown *Xcv* strains. GFP fluorescence of MMA-grown *Xcv* strains 85-10 (wt) and Δ*sX13* carrying pB or p*sX13* and carrying GFP-reporter plasmids pFX*3927*, pFX*3927*MUT, pFX*pilH* or pFX*pilH*MUT. pFX*3927*MUT and pFX*pilH*MUT contain a mutated 4G- and 5G-motif, respectively. *Xcv* autofluorescence was determined using pFX0. GFP fluorescence of the wt was set to 1. Data points and error bars represent mean values and standard deviations obtained from three independent experiments. Statistically significant differences compared to the wt are indicated by an asterisk (*t*-test; *P*<0.03). For comparison, see Figure 8A and C.
(EPS)

Figure S6 mRNA amount of *XCV3927::gfp* is sX13-independent. The *XCV3927::gfp* mRNA amount in NYG-grown *Xcv* strains 85-10 (wt) and Δ*sX13* carrying pB, p*sX13* or mutated sX13-derivatives and containing pFX*3927* or pFX*3927*MUT was analyzed by qRT-PCR using *gfp*-specific oligonucleotides. The RNA level in the wt was set to 1. Data points and error bars represent mean values and standard deviations obtained with three independent biological samples. For comparison, see Figure 7B and Figure 8A.
(EPS)

Figure S7 sX13 posttranscriptionally affects XCV3927::GFP and PilH::GFP synthesis. GFP fluorescence of NYG-grown *Xcv* strains 85-10 (wt), Δ*sX13* and Δ*sX13* containing chromosomally re-integrated *sX13* (Δ*sX13*+*sX13*ch); strains express *XCV3927::gfp* (pFXpl-3927) or *pilH::gfp* (pFXpl-*pilH*) under control of p*lac*. pFXMUT derivatives contain a mutated 4G-motif. *Xcv* autofluorescence was determined using pFX0 and is indicated by dashed line. GFP fluorescence of the wt carrying pFXpl-*3927* or pFXpl-*pilH* was set to 1. Data points and error bars represent mean values and standard deviations obtained from

three independent experiments. Asterisks indicate statistically significant differences (*t*-test; $P<0.03$).
(EPS)

Figure S8   Translation of HrpG::GFP and HrpX::GFP is sX13-independent. GFP fluorescence of NYG-grown *Xcv* strains 85-10 (wt) and Δ*sX13* expressing *hrpG::gfp* (pFXpl-*hrpG*) or *hrpX::gfp* (pFXpl-*hrpX*) under control of p*lac*. *Xcv* autofluorescence was determined using pFX0 and is indicated by dashed line. GFP fluorescence of the wt was set to 1. Data points and error bars represent mean values and standard deviations obtained from three independent experiments. Differences were not statistically significant (*t*-test; $P<0.03$).
(EPS)

Table S1   Bacterial strains, plasmids and oligonucleotides used in this study.
(PDF)

Table S2   sX13-regulated genes identified by microarray and qRT-PCR analysis.
(PDF)

Author Contributions

Conceived and designed the experiments: CS UA JB JS AB UB. Performed the experiments: CS UA JB JS. Analyzed the data: CS UA JB UB. Contributed reagents/materials/analysis tools: AB. Wrote the paper: CS.

References

1. Storz G, Vogel J, Wassarman KM (2011) Regulation by small RNAs in bacteria: expanding frontiers. Mol Cell 43: 880–891.
2. Richards GR, Vanderpool CK (2011) Molecular call and response: the physiology of bacterial small RNAs. Biochim Biophys Acta 1809: 525–531.
3. Waters LS, Storz G (2009) Regulatory RNAs in bacteria. Cell 136: 615–628.
4. Pfeiffer V, Papenfort K, Lucchini S, Hinton JC, Vogel J (2009) Coding sequence targeting by MicC RNA reveals bacterial mRNA silencing downstream of translational initiation. Nat Struct Mol Biol 16: 840–846.
5. Papenfort K, Vogel J (2010) Regulatory RNA in bacterial pathogens. Cell Host Microbe 8: 116–127.
6. Gottesman S, Storz G (2011) Bacterial small RNA regulators: versatile roles and rapidly evolving variations. Cold Spring Harbor Perspectives in Biology 3: a003798.
7. Papenfort K, Said N, Welsink T, Lucchini S, Hinton JC, et al. (2009) Specific and pleiotropic patterns of mRNA regulation by ArcZ, a conserved, Hfq-dependent small RNA. Mol Microbiol 74: 139–158.
8. Majdalani N, Cunning C, Sledjeski D, Elliott T, Gottesman S (1998) DsrA RNA regulates translation of RpoS message by an anti-antisense mechanism, independent of its action as an antisilencer of transcription. Proc Natl Acad Sci U S A 95: 12462–12467.
9. Soper T, Mandin P, Majdalani N, Gottesman S, Woodson SA (2010) Positive regulation by small RNAs and the role of Hfq. Proc Natl Acad Sci U S A 107: 9602–9607.
10. Göpel Y, Görke B (2012) Rewiring two-component signal transduction with small RNAs. Curr Opin Microbiol 15: 132–139.
11. Vogel J, Luisi BF (2011) Hfq and its constellation of RNA. Nat Rev Microbiol 9: 578–589.
12. Chao Y, Vogel J (2010) The role of Hfq in bacterial pathogens. Curr Opin Microbiol 13: 24–33.
13. Wilms I, Overlöper A, Nowrousian M, Sharma CM, Narberhaus F (2012) Deep sequencing uncovers numerous small RNAs on all four replicons of the plant pathogen *Agrobacterium tumefaciens*. RNA Biol 9: 446–457.
14. Filiatrault MJ, Stodghill PV, Bronstein PA, Moll S, Lindeberg M, et al. (2010) Transcriptome analysis of *Pseudomonas syringae* identifies new genes, noncoding RNAs, and antisense activity. J Bacteriol 192: 2359–2372.
15. Findeiß S, Schmidtke C, Stadler PF, Bonas U (2010) A novel family of plasmid-transferred anti-sense ncRNAs. RNA Biol 7: 120–124.
16. Schmidtke C, Findeiß S, Sharma CM, Kuhfuss J, Hoffmann S, et al. (2012) Genome-wide transcriptome analysis of the plant pathogen *Xanthomonas* identifies sRNAs with putative virulence functions. Nucleic Acids Res 40: 2020–2031.
17. Jiang RP, Tang DJ, Chen XL, He YQ, Feng JX, et al. (2010) Identification of four novel small non-coding RNAs from *Xanthomonas campestris* pathovar *campestris*. BMC Genomics 11: 10.1186/1471-2164-1111-1316.
18. Liang H, Zhao YT, Zhang JQ, Wang XJ, Fang RX, et al. (2011) Identification and functional characterization of small non-coding RNAs in *Xanthomonas oryzae* pathovar *oryzae*. BMC Genomics 12: 10.1186/1471-2164-1112-1187.
19. Chai Y, Winans SC (2005) A small antisense RNA downregulates expression of an essential replicase protein of an *Agrobacterium tumefaciens* Ti plasmid. Mol Microbiol 56: 1574–1585.
20. Wilms I, Voss B, Hess WR, Leichert LI, Narberhaus F (2011) Small RNA-mediated control of the *Agrobacterium tumefaciens* GABA binding protein. Mol Microbiol 80: 492–506.
21. Cui Y, Chatterjee A, Liu Y, Dumenyo CK, Chatterjee AK (1995) Identification of a global repressor gene, *rsmA*, of *Erwinia carotovora* subsp. carotovora that controls extracellular enzymes, N-(3-oxohexanoyl)-L-homoserine lactone, and pathogenicity in soft-rotting *Erwinia* spp. J Bacteriol 177: 5108–5115.

22. Liu Y, Cui Y, Mukherjee A, Chatterjee AK (1998) Characterization of a novel RNA regulator of *Erwinia carotovora* ssp. *carotovora* that controls production of extracellular enzymes and secondary metabolites. Mol Microbiol 29: 219–234.
23. Cui Y, Chatterjee A, Yang H, Chatterjee AK (2008) Regulatory network controlling extracellular proteins in *Erwinia carotovora* subsp. carotovora: FlhDC, the master regulator of flagellar genes, activates *rsmB* regulatory RNA production by affecting *gacA* and *hexA* (*lrhA*) expression. J Bacteriol 190: 4610–4623.
24. Leyns F, De Cleene M, Swings JG, De Ley J (1984) The host range of the genus *Xanthomonas*. The Botanical Review 50: 308–356.
25. Chan JW, Goodwin PH (1999) The molecular genetics of virulence of *Xanthomonas campestris*. Biotechnol Adv 17: 489–508.
26. Mhedbi-Hajri N, Darrasse A, Pigne S, Durand K, Fouteau S, et al. (2011) Sensing and adhesion are adaptive functions in the plant pathogenic xanthomonads. BMC Evol Biol 11: 10.1186/1471-2148-1111-1167.
27. Tang X, Xiao Y, Zhou JM (2006) Regulation of the type III secretion system in phytopathogenic bacteria. Mol Plant Microbe Interact 19: 1159–1166.
28. Ghosh P (2004) Process of protein transport by the type III secretion system. Microbiol Mol Biol Rev 68: 771–795.
29. Büttner D, Bonas U (2010) Regulation and secretion of *Xanthomonas* virulence factors. FEMS Microbiol Rev 34: 107–133.
30. Fouhy Y, Lucey JF, Ryan RP, Dow JM (2006) Cell-cell signaling, cyclic di-GMP turnover and regulation of virulence in *Xanthomonas campestris*. Res Microbiol 157: 899–904.
31. Jones JB, Stall RE, Bouzar H (1998) Diversity among xanthomonads pathogenic on pepper and tomato. Annu Rev Phytopathol 36: 41–58.
32. Thieme F, Koebnik R, Bekel T, Berger C, Boch J, et al. (2005) Insights into genome plasticity and pathogenicity of the plant pathogenic bacterium *Xanthomonas campestris* pv. vesicatoria revealed by the complete genome sequence. J Bacteriol 187: 7254–7266.
33. Bonas U, Schulte R, Fenselau S, Minsavage GV, Staskawicz BJ, et al. (1991) Isolation of a gene-cluster from *Xanthomonas campestris* pv. *vesicatoria* that determines pathogenicity and the hypersensitive response on pepper and tomato. Mol Plant Microbe Interact 4: 81–88.
34. White FF, Potnis N, Jones JB, Koebnik R (2009) The type III effectors of *Xanthomonas*. Mol Plant Pathol 10: 749–766.
35. Klement Z (1982) Hypersensitivity. In: Mount MS, Lacy GH, editors. Phytopathogenic Prokaryotes. New York: Academic Press. pp. 149–177.
36. White FF, Yang B, Johnson LB (2000) Prospects for understanding avirulence gene function. Curr Opin Plant Biol 3: 291–298.
37. Schulte R, Bonas U (1992) Expression of the *Xanthomonas campestris* pv. vesicatoria *hrp* gene cluster, which determines pathogenicity and hypersensitivity on pepper and tomato, is plant inducible. J Bacteriol 174: 815–823.
38. Wengelnik K, Marie C, Russel M, Bonas U (1996) Expression and localization of HrpA1, a protein of *Xanthomonas campestris* pv. vesicatoria essential for pathogenicity and induction of the hypersensitive reaction. J Bacteriol 178: 1061–1069.
39. Wengelnik K, Van den Ackerveken G, Bonas U (1996) HrpG, a key *hrp* regulatory protein of *Xanthomonas campestris* pv. *vesicatoria* is homologous to two-component response regulators. Mol Plant Microbe Interact 9: 704–712.
40. Wengelnik K, Bonas U (1996) HrpXv, an AraC-type regulator, activates expression of five of the six loci in the *hrp* cluster of *Xanthomonas campestris* pv. vesicatoria. J Bacteriol 178: 3462–3469.
41. Wengelnik K, Rossier O, Bonas U (1999) Mutations in the regulatory gene *hrpG* of *Xanthomonas campestris* pv. vesicatoria result in constitutive expression of all *hrp* genes. J Bacteriol 181: 6828–6831.

42. Noël L, Thieme F, Nennstiel D, Bonas U (2001) cDNA-AFLP analysis unravels a genome-wide *hrpG*-regulon in the plant pathogen *Xanthomonas campestris* pv. *vesicatoria*. Mol Microbiol 41: 1271–1281.

43. Guo Y, Figueiredo F, Jones J, Wang N (2010) HrpG and HrpX play global roles in coordinating different virulence traits of *Xanthomonas axonopodis* pv. *citri*. Mol Plant Microbe Interact 24: 649–661.

44. Ronald PC, Staskawicz BJ (1988) The avirulence gene *avrBs1* from *Xanthomonas campestris* pv. *vesicatoria* encodes a 50-kD protein. Mol Plant Microbe Interact 1: 191–198.

45. Escolar L, Van Den Ackerveken G, Pieplow S, Rossier O, Bonas U (2001) Type III secretion and *in planta* recognition of the *Xanthomonas* avirulence proteins AvrBs1 and AvrBsT. Mol Plant Pathol 2: 287–296.

46. Beisel CL, Storz G (2010) Base pairing small RNAs and their roles in global regulatory networks. FEMS Microbiol Rev 34: 866–882.

47. Møller T, Franch T, Udesen C, Gerdes K, Valentin-Hansen P (2002) Spot 42 RNA mediates discoordinate expression of the *E. coli* galactose operon. Genes Dev 16: 1696–1706.

48. Beisel CL, Storz G (2011) The base-pairing RNA spot 42 participates in a multioutput feedforward loop to help enact catabolite repression in *Escherichia coli*. Mol Cell 41: 286–297.

49. Dondrup M, Albaum SP, Griebel T, Henckel K, Jünemann S, et al. (2009) EMMA 2–a MAGE-compliant system for the collaborative analysis and integration of microarray data. BMC Bioinformatics 10: 50.

50. Mhedbi-Hajri N, Jacques MA, Koebnik R (2011) Adhesion mechanisms of plant-pathogenic *Xanthomonadaceae*. Adv Exp Med Biol 715: 71–89.

51. Jarrell KF, McBride MJ (2008) The surprisingly diverse ways that prokaryotes move. Nat Rev Microbiol 6: 466–476.

52. Schulze S, Kay S, Büttner D, Egler M, Eschen-Lippold L, et al. (2012) Analysis of new type III effectors from *Xanthomonas* uncovers XopB and XopS as suppressors of plant immunity. New Phytol 195: 894–911.

53. Urban JH, Vogel J (2007) Translational control and target recognition by *Escherichia coli* small RNAs *in vivo*. Nucleic Acids Res 35: 1018–1037.

54. Engler C, Kandzia R, Marillonnet S (2008) A one pot, one step, precision cloning method with high throughput capability. PLoS One 3: e3647.

55. Zuker M (2003) Mfold web server for nucleic acid folding and hybridization prediction. Nucleic Acids Res 31: 3406–3415.

56. Bailey TL (2011) DREME: motif discovery in transcription factor ChIP-seq data. Bioinformatics 27: 1653–1659.

57. Hengge R (2009) Principles of c-di-GMP signalling in bacteria. Nat Rev Microbiol 7: 263–273.

58. Ryan RP, Fouhy Y, Lucey JF, Jiang BL, He YQ, et al. (2007) Cyclic di-GMP signalling in the virulence and environmental adaptation of *Xanthomonas campestris*. Mol Microbiol 63: 429–442.

59. Wilms I, Möller P, Stock AM, Gurski R, Lai EM, et al. (2012) Hfq influences multiple transport systems and virulence in the plant pathogen *Agrobacterium tumefaciens*. J Bacteriol 194: 5209–5217.

60. Barra-Bily L, Pandey SP, Trautwetter A, Blanco C, Walker GC (2010) The *Sinorhizobium meliloti* RNA chaperone Hfq mediates symbiosis of *S. meliloti* and alfalfa. J Bacteriol 192: 1710–1718.

61. Lenz DH, Mok KC, Lilley BN, Kulkarni RV, Wingreen NS, et al. (2004) The small RNA chaperone Hfq and multiple small RNAs control quorum sensing in *Vibrio harveyi* and *Vibrio cholerae*. Cell 118: 69–82.

62. Bardill JP, Zhao X, Hammer BK (2011) The *Vibrio cholerae* quorum sensing response is mediated by Hfq-dependent sRNA/mRNA base pairing interactions. Mol Microbiol 80: 1381–1394.

63. Novick RP, Ross HF, Projan SJ, Kornblum J, Kreiswirth B, et al. (1993) Synthesis of staphylococcal virulence factors is controlled by a regulatory RNA molecule. EMBO J 12: 3967–3975.

64. Geisinger E, Adhikari RP, Jin R, Ross HF, Novick RP (2006) Inhibition of *rot* translation by RNAIII, a key feature of *agr* function. Mol Microbiol 61: 1038–1048.

65. Boisset S, Geissmann T, Huntzinger E, Fechter P, Bendridi N, et al. (2007) *Staphylococcus aureus* RNAIII coordinately represses the synthesis of virulence factors and the transcription regulator Rot by an antisense mechanism. Genes Dev 21: 1353–1366.

66. Geissmann T, Chevalier C, Cros MJ, Boisset S, Fechter P, et al. (2009) A search for small noncoding RNAs in *Staphylococcus aureus* reveals a conserved sequence motif for regulation. Nucleic Acids Res 37: 7239–7257.

67. Benito Y, Kolb FA, Romby P, Lina G, Etienne J, et al. (2000) Probing the structure of RNAIII, the *Staphylococcus aureus agr* regulatory RNA, and identification of the RNA domain involved in repression of protein A expression. RNA 6: 668–679.

68. Argaman L, Altuvia S (2000) *fhlA* repression by OxyS RNA: kissing complex formation at two sites results in a stable antisense-target RNA complex. J Mol Biol 300: 1101–1112.

69. Sharma CM, Hoffmann S, Darfeuille F, Reignier J, Findeiss S, et al. (2010) The primary transcriptome of the major human pathogen *Helicobacter pylori*. Nature 464: 250–255.

70. Desnoyers G, Morissette A, Prevost K, Masse E (2009) Small RNA-induced differential degradation of the polycistronic mRNA iscRSUA. EMBO J 28: 1551–1561.

71. Nakagawa S, Niimura Y, Miura K, Gojobori T (2010) Dynamic evolution of translation initiation mechanisms in prokaryotes. Proc Natl Acad Sci U S A 107: 6382–6387.

72. Sharma CM, Darfeuille F, Plantinga TH, Vogel J (2007) A small RNA regulates multiple ABC transporter mRNAs by targeting C/A-rich elements inside and upstream of ribosome-binding sites. Genes Dev 21: 2804–2817.

73. Sharma CM, Papenfort K, Pernitzsch SR, Mollenkopf HJ, Hinton JC, et al. (2011) Pervasive post-transcriptional control of genes involved in amino acid metabolism by the Hfq-dependent GcvB small RNA. Mol Microbiol 81: 1144–1165.

74. Martin-Farmer J, Janssen GR (1999) A downstream CA repeat sequence increases translation from leadered and unleadered mRNA in *Escherichia coli*. Mol Microbiol 31: 1025–1038.

75. del Val C, Rivas E, Torres-Quesada O, Toro N, Jimenez-Zurdo JI (2007) Identification of differentially expressed small non-coding RNAs in the legume endosymbiont *Sinorhizobium meliloti* by comparative genomics. Mol Microbiol 66: 1080–1091.

76. Sahagan BG, Dahlberg JE (1979) A small, unstable RNA molecule of *Escherichia coli*: spot 42 RNA. I. Nucleotide sequence analysis. J Mol Biol 131: 573–592.

77. Rice PW, Dahlberg JE (1982) A gene between *polA* and *glnA* retards growth of *Escherichia coli* when present in multiple copies: physiological effects of the gene for spot 42 RNA. J Bacteriol 152: 1196–1210.

78. Joyce CM, Grindley ND (1982) Identification of two genes immediately downstream from the *polA* gene of *Escherichia coli*. J Bacteriol 152: 1211–1219.

79. Polayes DA, Rice PW, Garner MM, Dahlberg JE (1988) Cyclic AMP-cyclic AMP receptor protein as a repressor of transcription of the *spf* gene of *Escherichia coli*. J Bacteriol 170: 3110–3114.

80. del Val C, Romero-Zaliz R, Torres-Quesada O, Peregrina A, Toro N, et al. (2011) A survey of sRNA families in a-proteobacteria. RNA Biol 9: 119–129.

81. Daniels MJ, Barber CE, Turner PC, Sawczyc MK, Byrde RJW, et al. (1984) Cloning of genes involved in pathogenicity of *Xanthomonas campestris* pv. *campestris* using the broad host range cosmid pLAFR1. EMBO J 3: 3323–3328.

82. Ausubel FM, Brent R, Kingston RE, Moore DD, Seidman JG, et al. (1996) Current Protocols in Molecular Biology. New York: John Wiley & Sons.

83. Figurski DH, Helinski DR (1979) Replication of an origin-containing derivative of plasmid RK2 dependent on a plasmid function provided *in trans*. Proc Natl Acad Sci U S A 76: 1648–1652.

84. Szczesny R, Jordan M, Schramm C, Schulz S, Cogez V, et al. (2010) Functional characterization of the Xcs and Xps type II secretion systems from the plant pathogenic bacterium *Xanthomonas campestris* pv. *vesicatoria*. New Phytol 187: 983–1002.

85. Murillo J, Shen H, Gerhold D, Sharma A, Cooksey DA, et al. (1994) Characterization of pPT23B, the plasmid involved in syringolide production by *Pseudomonas syringae* pv. *tomato* PT23. Plasmid 31: 275–287.

86. Huguet E, Hahn K, Wengelnik K, Bonas U (1998) *hpaA* mutants of *Xanthomonas campestris* pv. *vesicatoria* are affected in pathogenicity but retain the ability to induce host-specific hypersensitive reaction. Mol Microbiol 29: 1379–1390.

87. Minsavage GV, Dahlbeck D, Whalen MC, Kearny B, Bonas U, et al. (1990) Gene-for-gene relationships specifying disease resistance in *Xanthomonas campestris* pv. *vesicatoria* - pepper interactions. Mol Plant Microbe Interact 3: 41–47.

88. Büttner D, Nennstiel D, Klüsener B, Bonas U (2002) Functional analysis of HrpF, a putative type III translocon protein from *Xanthomonas campestris* pv. *vesicatoria*. J Bacteriol 184: 2389–2398.

89. Rossier O, Van den Ackerveken G, Bonas U (2000) HrpB2 and HrpF from *Xanthomonas* are type III-secreted proteins and essential for pathogenicity and recognition by the host plant. Mol Microbiol 38: 828–838.

90. Waldminghaus T, Gaubig LC, Klinkert B, Narberhaus F (2009) The *Escherichia coli* ibpA thermometer is comprised of stable and unstable structural elements. RNA Biol 6: 455–463.

91. Hartmann RK, Bindereif A, Schön A, Westhof E (2005) Handbook of RNA biochemistry. Wiley-VCH, Weinheim, Germany 2: 636–637.

92. DeRisi JL, Iyer VR, Brown PO (1997) Exploring the metabolic and genetic control of gene expression on a genomic scale. Science 278: 680–686.

93. Mayer L, Vendruscolo CT, Silva WP, Vorhölter FJ, Becker A, et al. (2011) Insights into the genome of the xanthan-producing phytopathogen *Xanthomonas arboricola* pv. *pruni* 109 by comparative genomic hybridization. J Biotechnol 155: 40–49.

94. Serrania J, Vorhölter FJ, Niehaus K, Pühler A, Becker A (2008) Identification of *Xanthomonas campestris* pv. *campestris* galactose utilization genes from transcriptome data. J Biotechnol 135: 309–317.

95. Dudoit S, Yang YH, Callow MJ, Speed TP (2002) Statistical methods for identifying differentially expressed genes in replicated cDNA microarray experiments. Stat Sin 12: 111–140.

2.3.1.1. Anlagen zu Publikation 3

Die folgenden ‚Supporting Informations' zu Kapitel 2.3.1. enthalten die Abbildungen S1 bis S8 und die Tabellen S1 und S2.
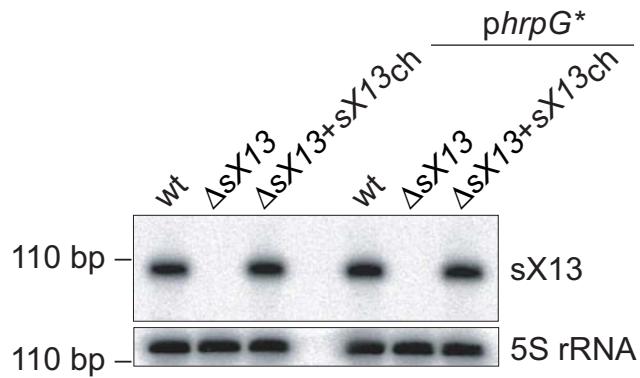
Figure S1. sX13 abundance is not affected by expression of HrpG*. *Xcv* 85-10 (wt), Δ*sX13* and Δ*sX13*+*sX13*ch and strains additionally expressing HrpG* were incubated for 3.5 hours in *hrp*-gene inducing medium XVM2 (see Figure 2B). Total RNA was analyzed by Northern blot using an sX13-specific probe. 5S rRNA was probed as loading control. The experiment was performed twice with similar results.
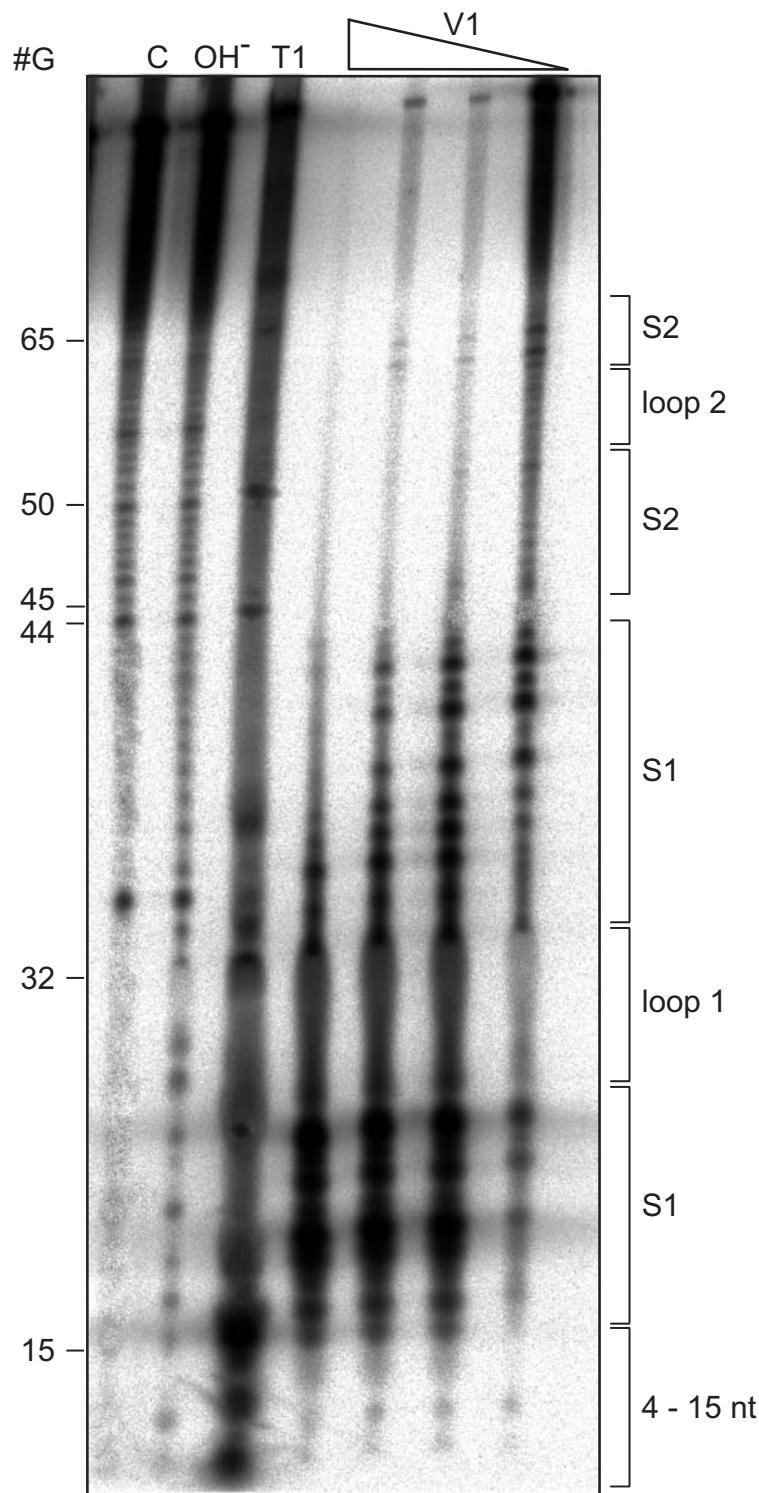
Figure S2. Structure probing of sX13. *In vitro* transcribed sX13 was 5'-labeled and treated with RNase T1 (T1) or alkaline hydroxyl (OH⁻) buffer for generation of nucleo-tide ladders and RNase V1 (V1) for mapping of base-paired regions. Lane 'C' contains untreated sX13; black triangle indicates decreasing concentrations of RNase V1; '#G' indicates positions of G residues; the deduced secondary structure is indicated on the right hand side (see Figure 6A).

Figure S3. Expression of sX13 derivatives. Total RNA of NYG-grown *Xcv* strains 85-10 (wt) and Δ*sX13* carrying pBRS (pB), p*sX13* or expressing mutated sX13-derivatives (see Figure 6) was analyzed by Northern blot using an sX13-specific probe. 5S rRNA was probed as loading control. The experiment was performed twice with similar results.

A
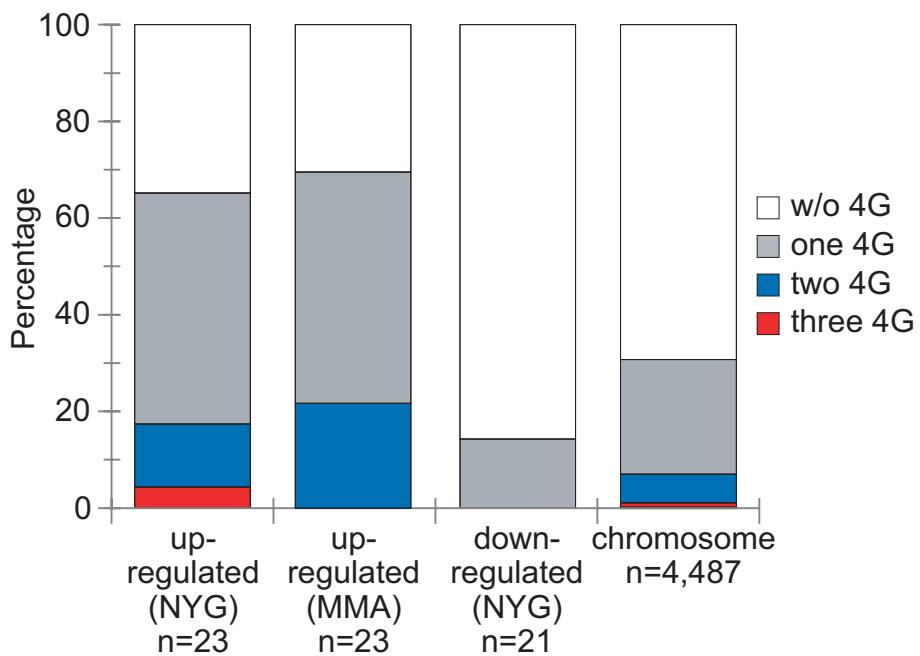
B
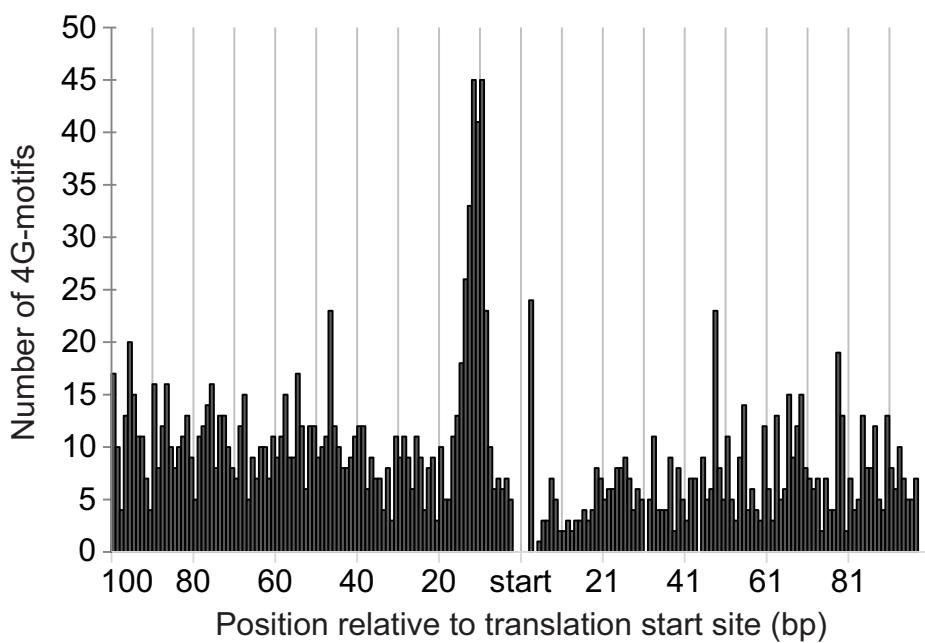
Figure S4. Distribution of 4G-motifs among sX13-regulated genes and chromosomally encoded *Xcv* genes. (A) Percentage of sX13-regulated genes identified by microarray analyses (see Table S2) and chromosomal CDSs in *Xcv* containing one or more 4G-motifs in region -100 to +100 relative to the TLS or in case of known TSSs, in the sequence comprising the 5'-UTR to position +100. The number of genes analyzed (n) is given below. (B) Distribution of 4G-motifs found in region -100 to +100 bp relative to the TLSs of 1,378 chromosomal CDSs [see (A)].

Figure S5. sX13-dependency of mRNA target::GFP synthesis in MMA-grown *Xcv* **strains.** GFP fluorescence of MMA-grown *Xcv* strains 85-10 (wt) and $\Delta sX13$ carrying pB or p*sX13* and carrying GFP-reporter plasmids pFX*3927*, pFX*3927*$_{MUT}$, pFX*pilH* or pFX*pilH*$_{MUT}$. pFX*3927*$_{MUT}$ and pFX*pilH*$_{MUT}$ contain a mutated 4G- and 5G-motif, respectively. *Xcv* autofluorescence was determined using pFX0. GFP fluorescence of the wt was set to 1. Data points and error bars represent mean values and standard deviations obtained from three independent experiments. Statistically significant differences compared to the wt are indicated by an asterisk (*t*-test; *P*<0.03). For comparison, see Figure 8A and C.

Figure S6. mRNA amount of *XCV3927::gfp* is sX13-independent.
The *XCV3927::gfp* mRNA amount in NYG-grown *Xcv* strains 85-10 (wt)
and Δ*sX13* carrying pB, p*sX13* or mutated sX13-derivatives and contai-
ning pFX*3927* or pFX*3927*$_{MUT}$ was analyzed by qRT-PCR using *gfp*-
specific oligonucleotides. The RNA level in the wt was set to 1. Data
points and error bars represent mean values and standard deviations
obtained with three independent biological samples. For comparison,
see Figure 7B and Figure 8A.

Figure S7. sX13 posttranscriptionally affects XCV3927::GFP and PilH::GFP synthesis. GFP fluorescence of NYG-grown *Xcv* strains 85-10 (wt), Δ*sX13* and Δ*sX13* containing chromosomally re-integrated *sX13* (Δ*sX13*+*sX13*ch); strains express *XCV3927::gfp* (pFXpl-*3927*) or *pilH::gfp* (pFXpl-*pilH*) under control of p*lac*. pFX$_{MUT}$ derivatives contain a mutated 4G-motif. *Xcv* autofluorescence was determined using pFX0 and is indicated by the dashed line. GFP fluorescence of the wt carrying pFXpl-*3927* or pFXpl-*pilH* was set to 1. Data points and error bars represent mean values and standard deviations obtained from three independent experiments. Asterisks indicate statistically significant differences (*t*-test; *P*<0.03).

Figure S8. Translation of HrpG::GFP and HrpX::GFP is sX13-independent. GFP fluorescence of NYG-grown *Xcv* strains 85-10 (wt) and Δ*sX13* expressing *hrpG::gfp* (pFXpl-*hrpG*) or *hrpX::gfp* (pFXpl-*hrpX*) under control of p*lac*. *Xcv* autofluorescence was determined using pFX0 and is indicated by dashed line. GFP fluorescence of the wt was set to 1. Data points and error bars represent mean values and standard deviations obtained from three independent experiments. Differences were not statistically significant (*t*-test; *P*<0.03).

Table S1. Bacterial strains, plasmids and oligonucleotides used in this study.

| Strain or plasmid	Relevant characteristics[a]	Reference or source
Xanthomonas campestris **pv.** ***vesicatoria***		
85-10	Pepper-race 2; wild type; Rif[R]	[1]
Δ*sX13*	85-10 derivative deleted in *sX13*; Rif[R]	This study
Δ*sX13*+*sX13*ch	Δ*sX13* derivative containing re-integrated *sX13* at Δ*sX13* locus; Rif[R]	This study
hfq⁻	*hfq* frameshift mutant of strain 85-10; Rif[R]	This study
Δ*sX13hfq*⁻	85-10 derivative deleted in *sX13* and containing frameshift mutation in *hfq*; Rif[R]	This study
Escherichia coli		
DH5α λpir	F⁻ *recA hsdR17(rk⁻,mk⁺)* φ80d*lacZ* Δ*M15 [λpir]*	[2]
TOP10	F⁻ *mcr*A Δ(*mrr-hsd*RMS-*mcr*BC) φ80*lacZ*ΔM15 Δ*lac*X74 *rec*A1 *ara*D139 Δ(*ara-leu*) 7697 *gal*U *gal*K *rps*L (Str[R]) *end*A1 *nup*G	Invitrogen
Plasmids		
pRK2013	ColE1 replicon, TraRK⁺ Mob⁺; Km[R]	[3]
pFG72-1	Derivative of pUFR043 containing *hrpG**; Km[R]	[4]
pOK1	Suicide vector; *sacB sacQ mobRK2* oriR6K; Sm[R]	[5]
pOKΔ*sX13*	pOK1 derivative containing flanking regions of *sX13*; Sm[R]	This study
pOKint13	pOK1 derivative containing *sX13* locus; Sm[R]	This study
pOK-fs*hfq*	pOK1 derivative for frameshift mutation of *hfq*; Sm[R]	This study
pBRM-P	pBBR1MCS-5 derivative without promoter; Gm[R]	[6]
p*hfq*	pBRM-P derivative containing *hfq*; Gm[R]	This study
pBBR1mod1	pBBR1MCS-5 derivative without polylinker; Gm[R]	[6]
pBRS	pBBR1mod1 derivative for sRNA expression; Gm[R]	This study
p*sX13*	pBRS derivative expressing *sX13*; Gm[R]	This study
p*sX13*Δ5'	pBRS derivative expressing *sX13*Δ5 (lacks 14 nt at 5' end); Gm[R]	This study
p*L1*	p*sX13* derivative containing mutations in *sX13* loop 1; Gm[R]	This study
p*L2*	p*sX13* derivative containing mutations in *sX13* loop 2; Gm[R]	This study
p*L3*	p*sX13* derivative containing mutations in *sX13* loop 3; Gm[R]	This study
p*L1/2*	p*L2* derivative containing mutations in *sX13* loops 1 and 2; Gm[R]	This study
p*L1/3*	p*L1* derivative containing mutations in *sX13* loops 1 and 3; Gm[R]	This study
p*L2/3*	p*sX13* derivative containing mutations in *sX13* loops 2 and 3; Gm[R]	This study
pDSK602	Broad-host-range vector; contains triple *lacUV5* promoter; Sm[R]	[7]
pXG-1	GFP expression plasmid; Cm[R]	[8]
pFX-P	Golden Gate-compatible pDSK602 derivative without promoter for generation of translational *mRNA::gfp* fusions; Sm[R]	This study
pFX0	Promoterless pFX-P derivative; control plasmid for GFP reporter fusions (measurement of *Xcv* autofluorescence); Sm[R]	This study
pFX*3927*	pFX-P derivative for expression of XCV3927::GFP; Sm[R]	This study
pFX*hfq*	pFX-P derivative for expression of Hfq::GFP; Sm[R]	This study
pFX*pilH*	pFX-P derivative for expression of PilH::GFP; Sm[R]	This study
pFX*0612*	pFX-P derivative for expression of XCV0612::GFP; Sm[R]	This study
pFX*3927*MUT	pFX-P derivative containing mutation in 4G-motif within 5'-UTR of *XCV3927*; Sm[R]	This study
pFX*hfq*MUT	pFX-P derivative containing mutation in 4G-motif within CDS of *hfq*; Sm[R]	This study
pFX*pilH*MUT	pFX-P derivative containing mutation in 5G-motif within 5'-UTR of *pilH*; Sm[R]	This study
pFXpl-*3927*	pFX-P derivative for p*lac*-driven expression of XCV3927::GFP; Sm[R]	This study
pFXpl-*pilH*	pFX-P derivative for p*lac*-driven expression of PilH::GFP; Sm[R]	This study
pFXpl-*3927*MUT	pFX-P derivative containing p*lac* and mutation in 4G-motif within 5'-UTR of *XCV3927*; Sm[R]	This study
pFXpl-*pilH*MUT	pFX-P derivative containing p*lac* and mutation in 5G-motif within 5'-UTR of *pilH*; Sm[R]	This study
pFXpl-*hrpG*	pFX-P derivative for p*lac*-driven expression of HrpG::GFP; Sm[R]	This study
pFXpl-*hrpX*	pFX-P derivative for p*lac*-driven expression of HrpX::GFP; Sm[R]	This study
pUC-13T7	pUC57 derivative containing T7 promoter upstream of *sX13*; Ap[R]; Sm[R]	This study

Oligonucleotide	Sequence[b]	Purpose
pBRS-EcoRI-fw	AAC<u>CTTAAG</u>ATTCCACACAACATACGAGC	Generation of pBRS
pBRS-NcoI-rev	CGT<u>CCATGG</u>GCAAATATTATA	Generation of pBRS
sX13-fw	TCA<u>GAATTC</u>GCGCAACGCCTGTCGGTAGA	Generation of p*sX13*
sX13-rev	GCT<u>AAGCTT</u>GCGCATAGTGGAAGGACACAAAT	Generation of p*sX13*
sX13d5-fw	TGGGAATTCGATCTCTCCCATCCCCTGG	Generation of p*sX13*Δ5

Oligonucleotide	Sequence[b]	Purpose
sX13d5-rev	TGGAAGCTTATAAAAAGCCCCGCAGACCAG	Generation of ps*X13*Δ5
L1-fw	CGGAAACTCCTCCCCAAGTTT	Generation of p*L1*
L1-rev	CTCCGAGATCTGCTCCAGCGCATGGGAG	Generation of p*L1*
L2-fw	AGCGGAAACTCCTGCGCAAGTTTCCGTTCC	Generation of p*L2*
L2-rev	CCGAGATCTGCTCCAGGGGATG	Generation of p*L2*
L3-fw	CCCCGCCGACCTGCGCCTGGTCTGC	Generation of p*L3*
L3-rev	CCAGGGAACGGAAACTTGGGGA	Generation of p*L3*
L1/2-rev	CCGAGATCTGCTCCAGCGCATGGGAGAGATC	Generation of p*L1/2*
L2/3-rev	CCAGGGAACGGAAACTTGCGCAGGAGTTTCC	Generation of p*L2/3*
plac-fw	TTTGGTCTCTATTCTGAGCGCAACGCAATTAATG	Generation of pFXpl
plac-rev	TTTGGTCTCTCCACCCACACAACATACGAGCCGG	Generation of pFXpl
pFX-lz-fw	GACATGCATGAATTCAGAGACCGCAGCTG	Generation of pFX-P
pFX-lz-rev	Phosphate-AGAGACCTTACAATTTCCATTCGC	Generation of pFX-P
pFXgfp-fw	Phosphate-GCTAGCAAAGGAGAAGAACTTTTCACTG	Generation of pFX-P
pFXgfp-rev	GACAGATCT AGCAAAACCCGTACCCTAGGTC	Generation of pFX-P
pFX0-fw	TTTGGTCTCTATTCCGCGAGGAAGAGGAAGAAGAA	Generation of pFX0
pFX0-rev	TTTGGTCTCTTAGCCATACAGCTACCCCAAAAGCGAAC	Generation of pFX0
pFX3927-fw	TTTGGTCTCTATTCCGGCAAGACGCTGTCATTCTAG	Generation of pFX*3927*
pFX3927-rev	TTTGGTCTCTTAGCAGCGACGACCGTACGAAGTC	Generation of pFX*3927*
pFXhfq-fw	TTTGGTCTCTATTCAGCGTGACCGCCATCAATTG	Generation of pFX*hfq*
pFXhfq-rev	TTTGGTCTCTTAGCATACACCGACACGGGCACC	Generation of pFX*hfq*
pFXpilH-fw	TTTGGTCTCTATTCACCCAGACGTGGTCGGAAC	Generation of pFX*pilH*
pFXpilH-rev	TTTGGTCTCTTAGCCCATTGACTGAAGACTGCCCTG	Generation of pFX*pilH*
pFX0612-fw	TTTGGTCTCTATTCATCGCGTGGTTTGTGATAAGTG	Generation of pFX*0612*
pFX0612-rev	TTTGGTCTCTTAGCCACCAGCGCTCTTAGTTGTTCTG	Generation of pFX*0612*
pFX3927mut-L-rev	TTTGGTCTCTGCGCAACAGGTCTGCGCACTATAGTCTAG	Generation of pFX*3927*$_{MUT}$
pFX3927mut-R-rev	TTTGGTCTCTGCGCAATCAGGCAAGAAGGCACCTATG	Generation of pFX*3927*$_{MUT}$
pFXhfqmut-L-rev	TTTGGTCTCTGCGCTTAGCCATCGAAAAATCCTCTTCA	Generation of pFX*hfq*$_{MUT}$
pFXhfqmut-R-rev	TTTGGTCTCTGCGCCAATCTTTACAGGACCCATTCCTC	Generation of pFX*hfq*$_{MUT}$
pFXpilHmut-L-rev	TTTGGTCTCTGCGCCTGGTCAGGCGTGGACGTAC	Generation of pFX*pilH*$_{MUT}$
pFXpilHmut-R-rev	TTTGGTCTCTGCGCAAAGGCAACATGGCTCGAATTATAT	Generation of pFX*pilH*$_{MUT}$
pFXpl3927-fw	TTTGGTCTCTGTGGACCTGTTGGGGAATCAGGCA	Generation of pFXpl-*3927*
pFXpl3927mut-fw	TTTGGTCTCTGTGGACCTGTTGCGCAATCAGGCA	Generation of pFXpl-*3927*$_{MUT}$
pFXplpilH-fw	TTTGGTCTCTGTGGGTTTCGTAGCGACGTCGGAAG	Generation of pFXpl-*pilH*
pFXpl-hrpG-fw	TTTGGTCTCTGTGGGTCCAGCTCCACTGGACTCTC	Generation of pFXpl-*hrpG*
pFXpl-hrpG-rev	TTTGGTCTCTTAGCGTCCTGCGTCAACAGGAACAC	Generation of pFXpl-*hrpG*
pFXpl-hrpX-fw	TTTGGTCTCTGTGGCGCCAGCGAGTTCGGCGC	Generation of pFXpl-*hrpX*
pFXpl-hrpX-rev	TTTGGTCTCTTAGCACGTTCTGCGTATGACAACGCA	Generation of pFXpl-*hrpX*
d13L-fw	CAGGATCCGCTGGGAGTACGGCTTCACG	Deletion of *sX13*
d13L-rev	AACAAGCTTATTTGTGTCCTTCCACTATGCGCA	Deletion of *sX13*
d13R-fw	AACAAGCTTATTGATGGATCGTGAAGATAACTG	Deletion of *sX13*
d13R-rev	GCTCTAGAAACTTCGGCCTGATGTACG	Deletion of *sX13*
int13L-fw	CAGGATCCCCGAGAGCATCCTGATGAGTTT	Δ*sX13* complementation
int13L-rev	TGCAACGTTAACAGCGATGCTGCAGGTG	Δ*sX13* complementation
int13R-fw	GTTAACGTTGCAGCGCTTGCGCATAGTG	Δ*sX13* complementation
int13R-rev	GCTCTAGAAGCTGATCGCCTGCGACTATT	Δ*sX13* complementation
hfqL-fw	TCAGGATCCAAATTGCCGATTCTGGCCGG	Mutation of *hfq*
hfqL-rev	TTTGGTCTCTCATTATGGGTCCTGTAAAGATTGCCC	Mutation of *hfq*
hfqR-fw	TTTGGTCTCTAATGCGCTGCGGCGCGAGC	Mutation of *hfq*
hfqR-rev	GCATCTAGAGCGTGGCGAACAATTGATCT	Mutation of *hfq*
seqhfq-fw	GAGCGTGACCGCCATCAATTG	Screening *hfq* mutation
seqhfq-rev	GAACTCCTCCATCACATCGTCTCG	Screening *hfq* mutation
pMphfq-fw	TTTGGTCTCTATTCAGCGTGACCGCCATCAATTG	Generation of p*hfq*
pMphfq-rev	CAGGGTCTCTCACCTTACTGCTCGACGTCGTCATCTTCCG	Generation of p*hfq*
sX13T7-fw	GAAATTAATACGACTCACTATAGGGCGCAACGCCTGTC	*in vitro* transcription
sX13T7-rev	TTATAAAAAGCCCCGCAGACCAGG	*in vitro* transcription
sX13-ITC-fw	TAATACGACTCACT	*in vitro* transcription
sX13-ITC-rev	ATAAAAAGCCCCGCA	*in vitro* transcription

Oligonucleotide	Sequence[b]	Purpose
q-16S-fw	TACGCTAATACCGCATACGAC	qRT-PCR
q-16S-rev	TGGCACGAAGTTAGCCGGTG	qRT-PCR
q-sX13-fw	CGCAACGCCTGTCGGTAGATCTC	qRT-PCR
q-sX13-rev	GGCCAGGGAACGGAAACTTG	qRT-PCR
q-gfp-fw	CCATGGCCAACACTTGTCACTA	qRT-PCR
q-gfp-rev	CAATGTTGTGGCGAATTTTGAA	qRT-PCR
q-algR-fw	ATCCGCAGGTCGAGGTGAT	qRT-PCR
q-algR-rev	ACCGGTTTCATCAGGTAATCCAG	qRT-PCR
q-asnB2-fw	GCCTACAACGGCGAGGTCTAT	qRT-PCR
q-asnB2-rev	ATCAGCTTGAAGGTGTGCTCGT	qRT-PCR
q-avrBs1-fw	AGGTCGCCACTCAGCAAGATAG	qRT-PCR
q-avrBs1-rev	TAGTCACCTCTTGGGGGTTTGA	qRT-PCR
q-cheY-fw	AGCGCACGTATCTTGGTGGT	qRT-PCR
q-cheY-rev	GCCTTGCCTTCGGATTTCTT	qRT-PCR
q-fliC-fw	GCACAGGTAATCAACACCAACG	qRT-PCR
q-fliC-rev	AGTTAGTTGCTTTCGCCGACTG	qRT-PCR
q-hfq-fw	ATGGCTAAGGGGCAATCTTTACAGG	qRT-PCR
q-hfq-rev	CGTCGTCATCTTCCGCCTGAT	qRT-PCR
q-hrcJ-fw	ATCAGGTGTCCTATTCGCTGGA	qRT-PCR
q-hrcJ-rev	TTTCGTAGGTGAGTCCCTCCAC	qRT-PCR
q-hrpF-fw	CAAGTCGGAGCTTCAGATCGTT	qRT-PCR
q-hrpF-rev	CTCCAGTTCGGATTGATTGAGC	qRT-PCR
q-hrpG-fw	TCTCGACGTTTTCCGATGAA	qRT-PCR
q-hrpG-rev	CATCGCGGATCAGCTTGTAC	qRT-PCR
q-hrpX-fw	GATGAGGTCAGCTTGTTCGGTG	qRT-PCR
q-hrpX-rev	GTCTGTAAGGCCAACGTGCTCTG	qRT-PCR
q-pilE-fw	TAGCGGAGCGATTTCACACA	qRT-PCR
q-pilE-rev	TCTGACTGGAGCCCCTTTGA	qRT-PCR
q-pilG-fw	AACTCGCAGGACTGAAGGTGAT	qRT-PCR
q-pilG-rev	GAACGGCTTGGTCAGATATTGC	qRT-PCR
q-pilH-fw	ACCGACAGGGCAGTCTTCAGT	qRT-PCR
q-pilH-rev	AACGCACCAATTGTCTGATGC	qRT-PCR
q-pilN-fw	TCTGACCACCGAGATCGACA	qRT-PCR
q-pilN-rev	GACCTTCCAGCGTGAGGATG	qRT-PCR
q-pilU-fw	AGAAGGCCTCGGACCTGTTC	qRT-PCR
q-pilU-rev	ACCATGCCCACCTGATTACG	qRT-PCR
q-xopJ-fw	CACGTCTTCTTACTCGGCCACT	qRT-PCR
q-xopJ-rev	CACGCTGGGAAACTACTGAGGT	qRT-PCR
q-xopS-fw	GCGATCATTCTGGAAGACCAGT	qRT-PCR
q-xopS-rev	TTGGCTTCAATCCTCGTCAGTT	qRT-PCR
q-XCV0173-fw	CCAAGCCCAAGGAGTTCTATTTCGAC	qRT-PCR
q-XCV0173-rev	CCAGGTGATGGCCCGGTACTG	qRT-PCR
q-XCV0612-fw	ATGAGGCGCATTTCTATGCTGT	qRT-PCR
q-XCV0612-rev	AATCAGTACGTCGAGGCGAATC	qRT-PCR
q-XCV2186-fw	AAGACTTTGTGCGCATTTCCAC	qRT-PCR
q-XCV2186-rev	AATCGGGTACTTCCTGCTTGGT	qRT-PCR
q-XCV2819-fw	TTGGATGCGTCAACGATCAG	qRT-PCR
q-XCV2819-rev	CGTCCATCGTCAACGTCGTA	qRT-PCR
q-XCV2821-fw	AGGACCGAGGCCATACGGAATG	qRT-PCR
q-XCV2821-rev	CGGCTGAAGAGTGACTTTCTGGTCTG	qRT-PCR
q-XCV3096-fw	TCCCTGCTACTGGCGTTGTT	qRT-PCR
q-XCV3096-rev	GATTCCGTTCGACCGTCTTG	qRT-PCR
q-XCV3572-fw	GTTGGTGCTGGTGAACGGTA	qRT-PCR
q-XCV3572-rev	ATATTCGGCAGACAGCGTCA	qRT-PCR
q-XCV3573-fw	GCAGCGACATTCACCTTGTG	qRT-PCR
q-XCV3573-rev	AGGTACAGCCGTGCACGCTG	qRT-PCR
q-XCV3927-fw	CCGCAAGGGCGATACCTTGTG	qRT-PCR
q-XCV3927-rev	CCATGTCCAGATCTTCGTTGTGCTTG	qRT-PCR

[a], Ap, ampicillin; Gm, gentamycin; Km, kanamycin; Rif, rifampicin; Sm, spectinomycin; [R], resistance. [b], Recognition sites of restriction enzymes are underlined.

References

1. Canteros BI (1990) Ph.D. thesis. University of Florida, Gainesville, FL.
2. Menard R, Sansonetti PJ, Parsot C (1993) Nonpolar mutagenesis of the *ipa* genes defines IpaB, IpaC, and IpaD as effectors of *Shigella flexneri* entry into epithelial cells. J Bacteriol 175: 5899-5906.
3. Figurski DH, Helinski DR (1979) Replication of an origin-containing derivative of plasmid RK2 dependent on a plasmid function provided *in trans*. Proc Natl Acad Sci U S A 76: 1648-1652.
4. Wengelnik K, Rossier O, Bonas U (1999) Mutations in the regulatory gene *hrpG* of *Xanthomonas campestris* pv. vesicatoria result in constitutive expression of all *hrp* genes. J Bacteriol 181: 6828-6831.
5. Huguet E, Hahn K, Wengelnik K, Bonas U (1998) *hpaA* mutants of *Xanthomonas campestris* pv. *vesicatoria* are affected in pathogenicity but retain the ability to induce host-specific hypersensitive reaction. Mol Microbiol 29: 1379-1390.
6. Szczesny R, Jordan M, Schramm C, Schulz S, Cogez V, et al. (2010) Functional characterization of the Xcs and Xps type II secretion systems from the plant pathogenic bacterium *Xanthomonas campestris* pv. *vesicatoria*. New Phytol 187: 983-1002.
7. Murillo J, Shen H, Gerhold D, Sharma A, Cooksey DA, et al. (1994) Characterization of pPT23B, the plasmid involved in syringolide production by *Pseudomonas syringae* pv. *tomato* PT23. Plasmid 31: 275-287.
8. Urban JH, Vogel J (2007) Translational control and target recognition by *Escherichia coli* small RNAs *in vivo*. Nucleic Acids Res 35: 1018-1037.

Table S2. sX13-regulated genes identified by microarray and qRT-PCR analysis.

Locus[a]	Annotated gene product[b]	4G-motif[c]	Microarray Fold-change ($\Delta sX13$ / wt)[d]		qRT-PCR Fold-change ($\Delta sX13$ / wt)[e]	
			NYG	MMA	NYG	MMA
Upregulated genes ($\Delta sX13$ / wt)						
XCV0678	AlgR; two-component system regulatory protein	a,a,a	1.8	—	2.5 ± 0.23	n.t.
XCV0730	Prc; tail-specific protease	—	—	1.6	n.t.	n.t.
XCV0950	conserved hypothetical protein	a,a	1.6	—	n.t.	n.t.
XCV1274	putative secreted protein	a	—	2.7	n.t.	n.t.
XCV1528	putative secreted protein	—	—	1.8	n.t.	n.t.
XCV1626	peptidyl-prolyl cis-trans isomerase	—	—	1.6	n.t.	n.t.
XCV1768	Hfq; host factor-I protein	b	1.6	—	2.4 ± 0.08	1.6 ± 0.31
XCV2041	putative signal transduction protein	a	—	1.9	n.t.	n.t.
XCV2185	conserved hypothetical protein	—	3.7	—	n.t.	n.t.
XCV2186	methyl-accepting chemotaxis protein	a	7.7	—	2.1 ± 0.34	10.2 ± 4.63
XCV2302	conserved hypothetical protein	—	1.5	1.6	n.t.	n.t.
XCV2341	conserved hypothetical protein	—	—	1.5	n.t.	n.t.
XCV2357	conserved hypothetical protein	—	1.7	—	n.t.	n.t.
XCV2499	putative membrane protein	a,b	2.0	—	n.t.	n.t.
XCV2565	conserved hypothetical protein	b	—	1.6	n.t.	n.t.
XCV2608	type IV secretion system VirJ-like protein	b	1.7	2.5	n.t.	n.t.
XCV2814	PilE; type IV pilus pilin	—	2.8	—	3.3 ± 0.36	n.t.
XCV2815	type IV pilus adhesin	b	2.1	—	n.t.	n.t.
XCV2819	type IV pilus assembly protein PilW	a	3.7	4.0	3.4 ± 0.37	5.5 ± 3.0
XCV2820	type IV pilus assembly protein PilV	a	2.9	—	n.t.	n.t.
XCV2821	type IV pilus assembly protein FimT	a	4.3	7.4	4.2 ± 0.32	3.4 ± 1.27
XCV2917	hypothetical protein	—	2.9	—	n.t.	n.t.
XCV3059	putative secreted protein	b	—	2.1	n.t.	n.t.
XCV3067	PilU; type IV pilus assembly protein ATPase	a	1.8	—	1.7 ± 0.29	n.t.
XCV3096	ComEA-related DNA uptake protein	—	—	4.2	n.t.	1.9 ± 0.12
XCV3151	hypothetical protein	b,b	—	1.6	n.t.	n.t.
XCV3230	PilJ; type IV pilus methyl-accepting chemotaxis protein	a	2.2	—	n.t.	n.t.
XCV3233	PilG; type IV pilus response regulator	a,b	—	2.0	2.3 ± 0.26	4.1 ± 1.71
XCV3353	PilA; type IV pilus assembly protein, major pilin	a	—	2.9	n.t.	n.t.
XCV3371	conserved hypothetical protein	a,b	—	1.8	n.t.	n.t.
XCV3376	hypothetical protein	—	2.0	—	n.t.	n.t.
XCV3497	PilQ; type IV pilus assembly protein	a,b	—	1.6	n.t.	n.t.
XCV3498	PilP; type IV pilus assembly protein	—	2.5	—	n.t.	n.t.
XCV3499	PilO; type IV pilus assembly protein	a	2.4	—	n.t.	n.t.
XCV3500	PilN; type IV pilus assembly protein	—	2.7	—	2.7 ± 0.16	n.t.
XCV3629	putative amidohydrolase family protein	a	1.9	—	n.t.	n.t.
XCV3727	conserved hypothetical protein	a	—	2.3	n.t.	n.t.
XCV3730	type IV pilus assembly protein	a,a	2.0	—	n.t.	n.t.
XCV3927	putative secreted protein	a	—	1.7	5.6 ± 0.45	8.3 ± 4.54
XCV4099	conserved hypothetical protein	b	—	1.7	n.t.	n.t.
XCV4117	conserved hypothetical protein	b	—	1.5	n.t.	n.t.
XCV4382	putative acetyltransferase	—	—	1.5	n.t.	n.t.
Downregulated genes ($\Delta sX13$ / wt)						
XCV0227	hypothetical protein	—	0.4	—	n.t.	n.t.
XCV0588	putative secreted protein; YceI-like family	—	0.5	—	n.t.	n.t.
XCV1188	conserved hypothetical protein	—	0.5	—	n.t.	n.t.
XCV1315	HrpX; AraC-type transcriptional regulator	—	0.6	—	0.6 ± 0.01	0.7 ± 0.13
XCV1787	predicted ATPase related to phosphate starvation-inducible protein PhoH	—	0.6	—	n.t.	n.t.
XCV1945	methyl-accepting chemotaxis protein	—	0.5	—	n.t.	n.t.
XCV1956	CheA1; chemotaxis protein	—	0.6	—	n.t.	n.t.
XCV1957	CheY; chemotaxis response regulator	—	0.4	—	0.1 ± 0.04	n.t.
XCV1958	putative anti-sigma factor antagonist	—	0.4	—	n.t.	n.t.
XCV2021	FliD; flagellar capping protein	—	0.7	—	n.t.	n.t.
XCV2022	FliC; flagellin and related hook-associated proteins	—	0.2	—	0.06 ± 0.03	1.0 ± 0.39
XCV2037	conserved hypothetical protein	—	0.3	—	n.t.	n.t.
XCV2276	hypothetical protein	b	0.6	—	n.t.	n.t.
XCV2282	conserved hypothetical protein	—	0.6	—	n.t.	n.t.
XCV2535	CydA; cytochrome D ubiquinol oxidase, subunit II	a	0.3	—	n.t.	n.t.
XCV2537	putative membrane protein	—	0.1	—	n.t.	n.t.
XCV3119	sigma-54 modulation protein	—	0.6	—	n.t.	n.t.
XCV3206	TonB-dependent outer membrane receptor	—	0.4	—	n.t.	n.t.
XCV3572	TonB-dependent outer membrane receptor	a	0.2	—	0.2 ± 0.04	0.9 ± 0.24
XCV4119	putative secreted protein	—	0.2	—	n.t.	n.t.
XCVd0093	putative secreted protein	—	0.5	—	n.t.	n.t.

Locus[a]	Annotated gene product[b]	4G-motif[c]	Microarray Fold-change ($\Delta sX13$ / wt)[d]		qRT-PCR Fold-change ($\Delta sX13$ / wt)[e]	
			NYG	MMA	NYG	MMA
Additional genes tested by qRT-PCR						
XCV0173	putative secreted protein	a,b,b,b	—	—	1.9 ± 0.19	0.8 ± 0.26
XCV0612	ATPase of the AAA+ class	a	—	—	1.0 ± 0.06	0.8 ± 0.26
XCV1533	AsnB2; asparagine synthase	b	—	—	1.0 ± 0.04	1.0 ± 0.17
XCV3232	PilH; type IV pilus response regulator	a	—	—	2.2 ± 0.07	1.9 ± 0.67
XCV3573	putative transcriptional regulator, AraC family	a	—	—	0.2 ± 0.11	n.t.
XCV0324	type III effector protein XopS	—	—	—	0.6 ± 0.05	n.t.

[a], bold letters indicate genes with known TSS [1]. [b], refers to Thieme *et al.* (2005) [2]. [c], presence of a 4G-motif within the 5'-UTR or 100 bp upstream of translation start codon if TSS is unknown (a) and within 100 bp downstream of the start codon (b) (see Figure S4). [d], genes not detected as expressed are marked with —. [e], values represent mean fold-change and standard deviation (see Figure 4); n.t. - not tested.

References

1. Schmidtke C, Findeiss S, Sharma CM, Kuhfuss J, Hoffmann S, et al. (2012) Genome-wide transcriptome analysis of the plant pathogen *Xanthomonas* identifies sRNAs with putative virulence functions. Nucleic Acids Res 40: 2020-2031.
2. Thieme F, Koebnik R, Bekel T, Berger C, Boch J, et al. (2005) Insights into genome plasticity and pathogenicity of the plant pathogenic bacterium *Xanthomonas campestris* pv. vesicatoria revealed by the complete genome sequence. J Bacteriol 187: 7254-7266.

2.3.1.2. Zusammenfassung der Ergebnisse

Der Artikel beschreibt die funktionelle Charakterisierung der *Xcv* sRNA sX13, welche in der vorangegangenen dRNA-Seq Analyse als HrpG-/ HrpX-unabhängig exprimierte und abundante sRNA identifiziert wurde (s. Kapitel 2.1.1.). Die Analyse einer *sX13* Deletionsmutante ergab, dass sX13 zur Virulenz von *Xcv* in suszeptiblen und der HR Induktion in resistenten Wirtspflanzen beiträgt. Während das *in planta* Wachstum der *sX13* Deletionsmutante mit dem des Wildtypstamms vergleichbar war, förderte sX13 das Wachstum von *Xcv* in Komplex- und Minimalmedium. Expressionsanalysen ergaben, dass sX13 die mRNA Akkumulation von *hrpX*, jedoch nicht von *hrpG*, beeinflusst und die Expression von Komponenten und Substraten des T3S Systems fördert. Da die ektopische Expression von HrpG*, einer konstitutiv aktiven Version von HrpG, die *sX13* Deletion kompensierte, liegt die Vermutung nahe, dass sX13 die Aktivität von HrpG beeinflusst. Mittels Northern Blot Analysen wurde nachgewiesen, dass sX13 in *Xcv* unter verschiedenen Stressbedingungen akkumuliert. Dies lässt eine Rolle von sX13 in der Anpassung von *Xcv* an sich verändernde Umweltbedingungen vermuten. ‚Microarray'- und quantitative ‚reverse transcription'-PCR (qRT-PCR) Analysen ergaben, dass sX13 die mRNA Akkumulation von 63 Genen beeinflusst, welche vermutlich an der Signaltransduktion, der Motilität, QS, transkriptioneller und posttranskriptioneller Regulation sowie der Virulenz beteiligt sind. Da sX13 die mRNA Akkumulation von *hfq* negativ beeinflusste, wurde untersucht, ob die verzögerte Virulenz der *sX13* Deletionsmutante auf einer De-Regulation der Hfq Synthese beruht. Die Inaktivierung von *hfq* hatte keinen Einfluss auf die Virulenz von *Xcv* oder die Akkumulation und Aktivität von sX13, was zeigt, dass sX13 Hfq-unabhängig agiert. Während 70% der mRNAs, deren Akkumulation durch sX13 gehemmt wurde, ‚G'-reiche Motive in der Umgebung des TLS aufweisen, ergaben Strukturanalysen von sX13 drei doppelsträngige Bereiche mit ‚C'-reichen apikalen Loops. Die Mutation der sX13 Loops beeinflusste in unterschiedlichem Maße die Virulenz von *Xcv* und die Akkumulation potentieller Ziel-mRNAs. Mittels eines GFP-Reportersystems konnte nachgewiesen werden, dass sX13 die Expression von *XCV3927*, *pilH* und *hfq* hemmt und dass die ‚C'-reichen sX13 Loops und ‚G'-reiche Motive in potentiellen Ziel-mRNAs für die sX13-vermittelte Repression der Proteinsynthese essentiell sind. Des Weiteren zeigen die Ergebnisse, dass ‚G'-reiche mRNA Motive in potentiellen sX13-Ziel mRNAs als translationale ‚enhancer' wirken können und in 5% aller proteinkodierenden *Xcv* Gene an der RBS Position lokalisiert sind.

2.4. Eigenanteil an den Publikationen

Publikation 1, Kapitel 2.1.1. und 2.1.1.1.

Schmidtke, C., Findeiß, S., Sharma, C.M., Kuhfuss, J., Hoffmann, S., Vogel, J., Stadler, P.F. and Bonas, U. (2012) Genome-wide transcriptome analysis of the plant pathogen *Xanthomonas* identifies sRNAs with putative virulence functions. *Nucleic Acids Res.*, **40**, 2020-2031.

Eigenanteil: Die Planung der Probenaufbereitung für die Pyrosequenzierung erfolgte in Zusammenarbeit mit J. Vogel, C. M. Sharma und U. Bonas. Die RNA Extraktion und DNase I-Behandlung wurde von mir durchgeführt. Die nachfolgende Behandlung der RNA und die Sequenzierung wurde von C. M. Sharma durchgeführt. Die bioinformatische Prozessierung der Sequenzierdaten, das ‚mapping‘ der ‚reads‘ sowie die Klassifizierung von TSSs wurde von S. Findeiß in Absprache mit mir durchgeführt. Die folgenden Analysen und Experimente habe ich selbstständig durchgeführt: manuelle Sichtung der Sequenzierdaten und Auswahl von ncRNA Kandidaten; RNA Präparation und Test der ncRNA Kandidaten mittels Northern Blot; 5‘-RACE Analyse von asX4 sowie 5‘- und 3‘-RACE Analyse von sX12; Deletionsmutagenese von *sX12*, Erstellung des Komplementationskonstrukts und Infektionsexperimente; Wachstumsanalysen von *Xcv* und *XcvΔsX12*; bioinformatische Analyse der 5‘-UTRs von Typ III Effektorgenen. Die 3‘-RACE Experimente von asX4 erfolgten in Zusammenarbeit mit J. Brock (geb. Kuhfuß). Die Ergebnisse des von S. Findeiß und S. Hoffmann entwickelten bioinformatischen Ansatzes zur automatischen TSS Identifizierung wurden stichprobenartig von mir überprüft. Die Abbildungen 2, 3, 4, S3 und S4A sowie die Tabellen 1 und S1 wurden von mir erstellt. Die Abbildung 1 und die Tabellen S2 bis S7 wurden in Zusammenarbeit mit S. Findeiß angefertigt. Abbildung S4B wurde in Zusammenarbeit mit J. Brock erstellt. Die Anfertigung des Manuskripts erfolgte in Zusammenarbeit mit S. Findeiß und U. Bonas. Geteilte Erstautorenschaft mit S. Findeiß.

Publikation 2, Kapitel 2.2.1.

Findeiß, S., Schmidtke, C., Stadler, P.F. and Bonas, U. (2010) A novel family of plasmid-transferred anti-sense ncRNAs. *RNA Biol*, **7**, 120-124.

Eigenanteil: Identifizierung des *ptaRNA1* Lokus anhand von dRNA-Seq ‚reads‘ und bioinformatischen Analysen; RNA Präparation und Expressionsnachweis von PtaRNA1. Die Abbildung 1 und der zugehörige Abschnitt in ‚Materials and Methods‘ wurde von mir erstellt.

Publikation 3, Kapitel 2.3.1. und 2.3.1.1.

Schmidtke, C., Abendroth, U., Brock, J., Serrania, J., Becker, A. and Bonas, U. (2013) Small RNA sX13: a multifaceted regulator of virulence in the plant pathogen *Xanthomonas*. *PLoS Pathog.*, **9**, e1003626.

Eigenanteil: Die für die vorliegende Publikation generierten Plasmide und Stämme (Tabelle S1) wurden von mir geplant und erstellt, mit folgenden Ausnahmen: pOK-fs*hfq*, p*hfq* und die *Xcv* Stämme *hfq*⁻, *hfq*⁻(p*hfq*) und Δ*sX13hfq*⁻ wurden von J. Brock erstellt; p*hrpG** sowie die von mir generierten Plasmide pFX*3927* und pFX*hfq* wurden von U. Abendroth in *Xcv*Δ*sX13*+*sX13*$_{ch}$ bzw. in die Stämme *hfq*⁻ und Δ*sX13hfq*⁻ konjugiert; die *Xcv* Stämme Δ*sX13*(p*sX13*Δ5'), Δ*sX13*(p*L1*), Δ*sX13*(p*L2*), Δ*sX13*(p*L3*), Δ*sX13*(p*L1/2*), Δ*sX13*(p*L1/3*) und Δ*sX13*(p*L2/3*) sowie die entsprechenden Plasmide und pUC-13T7 wurden von U. Abendroth erstellt. Folgende Experimente und Analysen wurden selbstständig geplant, durchgeführt und ausgewertet: Untersuchung des bakteriellen Wachstums in Kultur und *in planta*; Infektionsstudien, mit Ausnahme der in Abbildung 5B und 6B dargestellten Experimente; Western und Northern Blot Experimente sowie quantitative RT-PCR Analysen, mit Ausnahme der in Abbildung 2B und 5A gezeigten Ergebnisse; Fluoreszenzmessungen mit Ausnahme des in Abbildung 5C gezeigten Experiments; RNA Präparation für ‚Microarray' Analysen. Die Probenvorbereitung, Hybridisierung und das Scannen der ‚Microarrays' wurde von J. Serrania durchgeführt, die Normalisierung und statistische Auswertung der Hybridisierungsergebnisse sowie alle weiteren statistischen und bioinformatischen Analysen erfolgten selbstständig. Die in Abbildung S2 dargestellte Strukturanalyse von sX13 wurde von U. Abendroth durchgeführt. Die Anfertigung der Abbildungen 2B, 5C und 6 bzw. 5A und 5B erfolgte in Zusammenarbeit mit U. Abendroth bzw. J. Brock. Alle weiteren Abbildungen und Tabellen wurden von mir erstellt. Das Manuskript wurde von mir in Zusammenarbeit mit U. Bonas angefertigt.

3. Diskussion

3.1. Das primäre Transkriptom von *Xcv*

Die Kenntnis von TSSs, transkribierten Sequenzregionen und ncRNAs ist die Grundlage für die Untersuchung posttranskriptioneller Regulationsmechanismen. Die dRNA-Seq Analyse von *Xcv* wurde im Jahr 2007 durchgeführt und ermöglichte erstmals Einblicke in das genomweite Transkriptom eines *Xanthomonas* Bakteriums. Hierbei wurden konservierte RNA Elemente, abundante asRNAs sowie sRNAs identifiziert, welche an der Wirt-Pathogen Interaktion beteiligt sind (s. Kapitel 2.1.1.; (188)). Für die dRNA-Seq Analyse wurde aus Kostengründen ein Gemisch von RNA der in NYG Komplexmedium angezogenen *Xcv* Stämme 85-10 und 85* in cDNA umgeschrieben und sequenziert (s. Kapitel 2.1.1.; Koop. mit C.M. Sharma; (188)). Der Stamm 85* exprimiert *hrpG** (s. Kapitel 1.4.2.) und wurde zur Detektion von Transkripten genutzt, die nur in Anwesenheit eines aktiven HrpG Proteins exprimiert werden, z.B. die ncRNAs sX12 und asX4. Allerdings erlauben die Sequenzierdaten nicht die Identifizierung differentiell exprimierter Gene, da die RNAs vor der RNA-Adaptorligation und cDNA Synthese gemischt wurde.

Im Rahmen dieser Arbeit wurden erstmals TSSs auf Basis eines statistischen Models identifiziert (s. Kapitel 2.1.1.; Koop. mit S. Findeiß und S. Hoffmann; (188)). Von den 1.421 in *Xcv* detektierten TSSs repräsentieren 831 wahrscheinlich die primären TSSs von 17,35% der 4.726 annotierten ORFs (s. Kapitel 2.1.1.; Koop. mit S. Findeiß; (188)). Für die meisten dieser TSSs wurden ‚reads' für die zugehörigen ORFs detektiert. Dagegen wiesen Gene, für die keine TSSs identifiziert wurden, keine bzw. nur sehr wenige ‚reads' auf (s. Kapitel 2.1.1.; Koop. mit S. Findeiß; (188)). Dies zeigt, dass durch die dRNA-Seq Analyse vor allem TSSs für hochabundante Transkripte detektiert wurden und deutet auf eine geringe Sensitivität der 454-Pyrosequenzierung hin. Insgesamt lieferte die Transkriptomanalyse nur eine geringe Anzahl auswertbarer ‚reads' (~90.000, tRNAs und rRNAs nicht inbegriffen), welche kaum Rückschlüsse auf Operonstrukturen erlauben (s. Kapitel 2.1.1.; Koop. mit S. Findeiß; (188)). Um auch TSSs für schwach exprimierte Gene zu identifizieren, sollten zukünftige Analysen Sequenzierplattformen mit größerer Sequenziertiefe nutzen, z.B. Illumina/ Solexa (10), und das Transkriptom von *Xcv* Stämmen untersuchen, welche unter verschiedenen Wachstumsbedingungen angezogen wurden.

dRNA-Seq Analysen anderer pflanzenpathogener Bakterien sind bislang auf *P. syringae* pv. *tomato* und *A. tumefaciens* beschränkt (53,113,258). Für *A. tumefaciens*, *S. enterica* ssp. *enterica* serovar Typhimurium (*S.* Typhimurium), *H. pylori* und das symbiotische Bakterium *Sinorhizobium meliloti* wurden mittels dRNA-Seq TSSs für 7,25%, 27,71%, 87,5% bzw. 60% der annotierten ORFs identifiziert (108,186,196,258). Die Ergebnisse für die verschiedenen Bakterienspezies sind allerdings nur eingeschränkt vergleichbar, da sich die Studien hinsichtlich der Anzahl und Art der gewählten

Wachstumsbedingungen, der Methode der RNA Präparation, den genutzten Sequenzierplattformen sowie den angewandten Kriterien zur Annotation von TSSs unterscheiden.

3.1.1. Diversität der 5'-UTR Längen von *Xcv* mRNAs

Die Mehrheit der mittels dRNA-Seq identifizierten 5'-UTRs in *Xcv* ist kürzer als 50 Nt und damit vergleichbar zu 5'-UTR Längen, die für *S.* Typhimurium und *H. pylori* beschrieben wurden (s. Kapitel 2.1.1.; Koop. mit S. Findeiß; (188))(108,196,204). Anders als in diesen Bakterien und *E. coli* weisen mRNAs in *Xcv* keine konservierte SD-Sequenz auf, wenngleich die Anti-SD Sequenz der 16S rRNA (s. Kapitel 1.3.1.) in *E. coli* und *Xcv* identisch ist (148,200). Aufgrund des hohen G+C Gehalts von *Xanthomonas* Genomen kommen SD-ähnliche (d.h. ‚G'-reiche) Sequenzmotive ähnlich häufig an RBS Positionen wie in anderen genomischen Regionen vor (148). Eine SD-abhängige Initiation der Translation in *Xanthomonas* ist daher insgesamt unwahrscheinlich. Dies lässt alternative Mechanismen der Translationsinitiation vermuten, z.B. durch das ribosomale Protein S1 (RPS1). In *E. coli* vermittelt RPS1 die Translationsinitiation unabhängig von der Gegenwart der SD-Sequenz (17,74,106). RPS1 ist in Gram-negativen Bakterien essentiell für die Translation von 5'-UTR-enthaltenden mRNAs, wird jedoch nicht für die Translation von ‚leaderless' mRNAs (lmRNAs), d.h. von mRNAs mit 5'-UTR Längen von weniger als 10 Nt, benötigt (74,176,185).

Zu Beginn dieser Arbeit waren nur wenige bakterielle lmRNAs bekannt und wurden deshalb als Ausnahmefälle betrachtet (140). Überraschenderweise wurden allein in *Xcv* TSSs für 118 lmRNAs detektiert, welche 14% der mRNAs entsprechen, für die in dieser Arbeit TSSs identifiziert wurden (s. Kapitel 2.1.1.; Koop. mit S. Findeiß; (188)). Dagegen machen lmRNAs in *S.* Typhimurium und *H. pylori* lediglich 1,2% bzw. 2,2% der transkribierten mRNAs aus (108,196). Während die Initiation der Translation von 5'-UTR-enthaltenden mRNAs durch die Bindung von 30S Ribosomenuntereinheiten an die RBS vermittelt wird (s. Kapitel 1.3.1.), wurde für lmRNAs in *E. coli* nachgewiesen, dass assemblierte 70S Ribosomen in Gegenwart der fMET-tRNA^{fMet} direkt an das ‚AUG' TLS der lmRNA binden und ohne Beteiligung von Translationsinitiationsfaktoren die Elongation der Translation vermitteln (21,153,223). In Übereinstimmung damit weisen 82% der in *Xcv* bekannten lmRNAs ein ‚AUG' als vorhergesagtes TLS auf (s. Kapitel 2.1.1.; Koop. mit S. Findeiß; (188)). Dennoch sollte in zukünftigen Analysen die Translation von *Xcv* lmRNAs nachgewiesen werden. Bislang ist ungeklärt, inwiefern die Koexistenz von 5'-UTR-enthaltenden mRNAs und lmRNAs zur Regulation der bakteriellen Genexpression beiträgt. Möglicherweise sind lmRNAs von posttranskriptioneller Regulation entkoppelt, da sRNAs und RNA-Bindeproteine überwiegend an 5'-UTRs binden. Hierfür spricht, dass einige *Xcv* lmRNAs Proteine mit vermutlich genereller zellulärer Funktion kodieren, z.B. die ß-Untereinheit der DNA Polymerase III und ribosomale Proteine (s. Kapitel 2.1.1.; (188)).

13% der in dieser Arbeit identifizierten 5'-UTRs von *Xcv* mRNAs weisen ungewöhnliche Längen von mehr als 150 Nt auf, einschließlich der 5'-UTRs der mRNAs, welche die Typ III Effektoren AvrBs1,

XopAA, XopB, XopC, XopD und XopN kodieren (s. Kapitel 2.1.1.; Koop. mit S. Findeiß; (188)). Für diese Effektoren wurden in *Xcv* Virulenzfunktionen beschrieben (24,82,191). Zukünftige Analysen könnten die 5'-UTRs der genannten *Xcv* Effektorgene hinsichtlich regulatorischer Funktionen untersuchen. Denkbar wäre, dass lange 5'-UTRs mRNAs stabilisieren, Riboswitches oder RNA Thermometer enthalten bzw. bislang unbekannte Proteine oder Peptide kodieren (218,261). Virulenzgene von *S. enterica* ssp. *enterica* serovar Typhi und *H. pylori* weisen ebenfalls lange 5'-UTRs auf, wobei die Funktion dieser Sequenzen bislang unbekannt ist (162,196).

Neben möglichen regulatorischen Funktionen könnten lange 5'-UTRs auch auf eine fehlerhafte Annotation von TLSs in *Xcv* hindeuten. In Übereinstimmung mit den dRNA-Seq Daten wurde kürzlich beschrieben, dass die kodierende Sequenz des Effektorgens *xopD* 215 Codons länger ist als bei der Genomannotation angenommen (s. Kapitel 2.1.1.; (188))(30,215). Die Ergebnisse der vorliegenden Arbeit zeigen außerdem, dass mindestens 71 *Xcv* ORFs kürzer sind als vermutet (s. Kapitel 2.1.1.; Koop. mit S. Findeiß; (188)). Basierend auf den dRNA-Seq Daten und der *Xcv* Genomsequenz könnten die TLSs von mehr als 10% der 4.726 ORFs fehlerhaft annotiert sein (215). Zudem wurde am Beispiel der sX6 RNA nachgewiesen, dass *Xcv* bislang unbekannte Proteine kodiert (s. Kapitel 2.1.1.; Koop. mit S. Findeiß und J. Brock; (188)). Insgesamt tragen die im Rahmen dieser Arbeit erhaltenen Ergebnisse maßgeblich zu einer ersten Verbesserung der *Xcv* Genomannotation bei. In zukünftigen Analysen könnten Genom-, Transkriptom- und Proteomstudien kombiniert werden, um TLS Vorhersagen experimentell zu bestätigen und bislang unbekannte ORFs zu identifizieren.

3.2. Konservierte RNAs mit vermutlich generellen zellulären Funktionen

Im Rahmen dieser Arbeit wurden mittels bioinformatischer Analysen der *Xcv* Genomsequenz fünf Gene vorhergesagt, die strukturell konservierte RNAs kodieren (RNase P, RtT RNA, SRP RNA, tmRNA und 6S RNA)(s. Kapitel 2.1.1.; Koop. mit S. Findeiß; (188)). Diese Transkripte standen nicht im Fokus dieser Arbeit, da in verschiedenen Bakterien generelle zelluläre Funktionen dieser RNAs, mit Ausnahme der RtT RNA, nachgewiesen wurden: Das Ribozym RNase P ist essentiell für die Prozessierung von Vorläufer-tRNAs, die SRP RNA (4.5S RNA) ist an der Membranlokalisierung von Proteinen mit Sec-Signalpeptiden beteiligt und tmRNA vermittelt die Termination der Translation im Falle der Unterbrechung des Translationsprozesses (s. Kapitel 1.3.2.)(49,72,81,142). Die genaue Funktion der RtT RNA, die in *E. coli* durch Prozessierung des *tyrT* tRNA Operons generiert wird, ist nicht bekannt (18). Da keine dRNA-Seq ‚reads' für den vorhergesagten *rtT* Lokus in *Xcv* detektiert wurden und die Region nicht Teil eines tRNA Operons ist, bleibt zu untersuchen, ob der Lokus eine ncRNA kodiert. Northern Blot Analysen ergaben eine Akkumulation der *Xcv* 6S RNA in der stationären Wachstumsphase (s. Kapitel 2.1.1.; (188)). Da zudem dRNA-Seq ‚reads' für eine pRNA detektiert wurden (s. Kapitel 2.1.1.; (188)), liegt der Schluss nahe, dass die 6S RNA in *Xcv* ähnliche

Funktionen wie in anderen Bakterien erfüllt und die Aktivität der RNA Polymerase reguliert (s. Kapitel 1.3.2.).

Mittels dRNA-Seq wurden TSSs von fünf potentiellen Riboswitches (TPP, SAM, SAH, Glycin und YybP-YkoY) in *Xcv* detektiert (s. Kapitel 2.1.1.; Koop. mit S. Findeiß; (188)). Zudem wurden drei Riboswitch Kandidaten (YybP-YkoY, FMN, Ado-Cbl) vorhergesagt, für die allerdings nur wenige bzw. keine Expressionsdaten vorliegen (s. Kapitel 2.1.1.; Koop. mit S. Findeiß; (188)). Mit Ausnahme der vorliegenden Arbeit wurden bislang keine Riboswitches in *Xanthomonas* spp. identifiziert. In *Xcv* und den meisten anderen Bakterien sind TPP, SAM, FMN bzw. Ado-Cbl Riboswitches mit Genen assoziiert, die an der Biosynthese von TPP, Methionin/ Cystein, FMN bzw. dem Transport von Ado-Cbl beteiligt sind (s. Kapitel 2.1.1.; (188)). Solche Riboswitches hemmen typischerweise die Expression der *cis*-lokalisierten Gene in Gegenwart der Liganden (50,69,136,137,147,231,232,260). Der SAH bzw. Glycin Riboswitch reguliert in *Xcv* vermutlich die Expression eines SAH-Hydrolase- bzw. Glycin-Dehydrogenase-kodierenden Gens (s. Kapitel 2.1.1.; (188)). In Gegenwart der Liganden fördern solche Riboswitches typischerweise die Synthese von Proteinen, die am Abbau von SAH bzw. Glycin beteiligt sind (129,194,239). Die zwei in *Xcv* identifizierten YybP-YkoY Riboswitch-Kandidaten sind, ähnlich wie in anderen Bakterien, mit Genen assoziiert, die potentielle Membranproteine kodieren (s. Kapitel 2.1.1.; (188))(6,134). Wenngleich die Liganden bislang unbekannt sind, wird angenommen, dass YybP-YkoY Elemente pH-Wert Änderungen perzipieren (134).

Da dRNA-Seq ‚reads' für die potentiellen TPP und SAM Riboswitches in *Xcv* ausschließlich für die Riboswitches, jedoch nicht für den jeweils flankierenden ORF detektiert wurden (s. Kapitel 2.1.1.; (188)), liegt der Schluss nahe, dass diese RNA Elemente die Expression der stromabwärts lokalisierten ORFs durch vorzeitige Termination der Transkription regulieren (s. Kapitel 1.3.1.). Eine ähnliche Funktionsweise wurde für TPP Riboswitches von *E. coli* und *Rhizobium etli* beschrieben (136,260). Interessanterweise wurden kürzlich im Gram-positiven Humanpathogen *L. monocytogenes* zwei SAM Riboswitches identifiziert, die auch in *trans* als sRNAs wirken können. Bei hohem zellulären SAM Level akkumulieren die Terminationsprodukte der Riboswitches (SreA und SreB; ~200 Nt) und hemmen die Synthese des Virulenzgenaktivators PrfA durch Basenpaarung mit der *prfA* mRNA (122). Dies wirft die Frage auf, ob in *Xcv* TPP und SAM Riboswitches die Virulenzgenexpression in Abhängigkeit von der Nährstoffverfügbarkeit beeinflussen können. Eine solche Funktion könnte durch Mutagenese der Riboswitches und Infektionsanalysen untersucht werden.

3.3. Identifizierung neuartiger ncRNAs in *Xcv*

Die dRNA-Seq Analyse von *Xcv* ergab antisense ‚reads' für 22% aller Nukleotide, die annotierten ORFs zugewiesen sind (s. Kapitel 2.1.1.; Koop. mit S. Findeiß; (188)). Vergleichbar mit diesem Ergebnis wurden asRNAs für 1-27% der proteinkodierenden Gene von *Mycoplasma pneumoniae*, *Synechocystis* sp. PCC 6803, *S. meliloti*, *A. tumefaciens*, *P. syringae* und *S. aureus* sowie für 46% der ORFs von *H. pylori* beschrieben (7,52,71,138,186,196,258).

In dieser Arbeit wurden TSSs für 178 abundante antisense Transkripte detektiert, von denen einige vermutlich die Aktivität mobiler Elemente oder die Expression von Virulenzgenen modulieren (s. Kapitel 2.1.1.; Koop. mit S. Findeiß; (188)). Beispielsweise wurden antisense ‚reads' für die meisten der 66 annotierten IS Elemente detektiert (188). Eine spezifische Zuordnung von dRNA-Seq ‚reads' zu IS Elementen ist jedoch nur eingeschränkt möglich, da beispielsweise Transposasegene von IS*1477* Elementen in 20 Kopien mit mehr als 90% Sequenzidentität vorliegen (215). Eine Funktion von asRNAs in der Regulation der Transposonaktivität wurde für *E. coli* beschrieben und wird für zahlreiche asRNAs in anderen Bakterien vermutet (14,124,186,258). Mit asRNAs assoziierte *Xcv* Pathogenitäts- bzw. Virulenzgene umfassen u.a. *hrcC* und die Typ III Effektorgene *avrBs1*, *xopAA*, *xopB*, *xopD*, *xopE2* und *xopO* (s. Kapitel 2.1.1.; (188)). *hrcC* kodiert das Sekretin des T3S Systems und ist essentiell für die Pathogenität von *Xcv* (16,253). Eine Virulenzfunktion wurde für XopAA, XopB und XopD beschrieben (102,143,191). Die genannten Gene werden transkriptionell durch HrpX induziert und weisen eine charakteristische PIP Box in den Promotorregionen auf (105,191,253). Dagegen ist das Expressionsmuster der hier identifizierten asRNAs unbekannt. Die Synthese von Typ III Effektoren kann durch Anzucht von *Xcv* in XVM2 Medium induziert werden, wohingegen die *in vitro* Sekretion nur in Minimalmedium A (pH 5,2) und in Gegenwart von HrpG* erfolgt (180,253). Denkbar wäre, dass die genannten asRNAs die Stabiliät und/ oder Translation der entsprechenden Effektor-mRNAs unter nicht-sekretorischen Bedingungen vermindern und dadurch die Akkumulation potentiell toxischer Proteine unterdrücken.

In dieser Arbeit wurden in *Xcv* 15 sRNAs (sX1-15), die 6S RNA und acht *cis*-kodierte asRNAs (asX1-7 und PtaRNA1) experimentell bestätigt (s. Kapitel 2.1.1.; (188)). Zudem ergaben Northern Blot Analysen für einige der 65 weiteren getesteten sRNA Kandidaten Signale für kurze Transkripte (C. Schmidtke und U. Bonas, unveröffentlicht). Allerdings korrelierten die detektierten RNA Längen nicht mit den dRNA-Seq Daten, da die entsprechenden Loci offenbar nicht vollständig durch ‚reads' abgedeckt wurden. Solche sRNA Kandidaten könnten in zukünftigen Analysen, z.B. durch Bestimmung der Transkriptenden, näher untersucht werden.

Für die meisten der verifizierten *Xcv* sRNAs und asRNAs ergaben Northern Blot Analysen mehrere Hybridisierungssignale, wobei die Abundanz einiger Signale in Abhängigkeit von HrpG und/ oder HrpX (sX4, asX1 und asX5) oder der Wachstumsphase (z.B. sX3 und 6S RNA) verändert war (s. Kapitel 2.1.1.; (188)). Die Größe von einigen der alternativen Hybridisierungssignale korrelierte mit

den Längen entsprechender dRNA-Seq ‚reads' (s. Kapitel 2.1.1.; (188)). Dennoch können Kreuzhybridisierungen der verwendeten Sonden mit anderen Transkripten nicht ausgeschlossen werden. Um die Spezifität der erhaltenen Hybridisierungssignale zu untersuchen, könnten andere Oligonukleotidsonden verwendet bzw. entsprechende sRNA Deletionsmutanten erstellt und mittels Northern Blot Analysen überprüft werden. Insgesamt deuten die Daten auf eine veränderte Stabilität bzw. Prozessierung von *Xcv* sRNAs unter verschiedenen Bedingungen hin (s. Kapitel 2.1.1.; (188)). Aufgrund der Rifampicin Resistenz der verwendeten *Xcv* Stämme war keine Untersuchung der RNA Stabilität mit Hilfe des üblicherweise verwendeten Transkriptionsinhibitors Rifampicin möglich. Die Prozessierung von *Xcv* sRNAs könnte durch Analyse der sRNA Expressionsmuster in RNase-Gen Mutanten untersucht werden. *Xcv* kodiert mehr als 15 vorhergesagte RNasen (215). Mögliche sRNA Prozessierungsprodukte wurden auch für *Xoo* und *A. tumefaciens* beschrieben (118,258). Zudem ist bekannt, dass die Aktivität enterobakterieller sRNAs, z.B. GlmZ, IstR-1 und MicX, durch RNase-vermittelte Prozessierung reguliert wird (43,98,233).

Phylogenetische Analysen ergaben, dass die in *Xcv* identifizierten sRNA und asRNA Gene, mit Ausnahme von *6S*, *sX8*, *asX6* und *ptaRNA1*, nur in den nahe verwandten Gattungen *Xanthomonas*, *Xylella* und *Stenotrophomonas* vorkommen (s. Kapitel 2.1.1.; Koop. mit S. Findeiß; (188)). Neben den im Rahmen dieser Arbeit identifizierten *Xcv* sRNAs wurden sRNAs bzw. Kandidaten für *Xcc* Stamm 8004, *Xoo* Stamm PXO99 und das opportunistische Humanpathogen *Stenotrophomonas maltophilia* Stamm K279a beschrieben (1,92,118,179). *S. maltophilia* gehört ebenfalls zur *Xanthomonadaceae* Familie der γ-Proteobakterien (37). Entgegen der Annahme der Autoren, dass die drei *Xcc* sRNAs *Xcc*2, *Xcc*3 und *Xcc*4 basenpaarende und *Xanthomonas*-spezifische sRNAs repräsentieren (92), handelt es sich dabei um konservierte Orthologe der 6S RNA, SRP RNA und 5S rRNA. Das *sRNA-Xcc1* Gen wird HrpX-abhängig transkribiert (36,92), kommt jedoch nicht in *Xcv* Stamm 85-10 vor. Kürzlich wurden acht weitere sRNAs und 16 sRNA Kandidaten in *Xcc* identifiziert, wobei der zugrundeliegende RNA-Seq Ansatz keine strangspezifische Zuordnung der Sequenzierdaten erlaubt (1). Interessanterweise wurde eine verminderte Virulenz für einen *Xcc* Stamm beschrieben, welcher Deletionen in drei Rpf-/ DSF-abhängig exprimierten sRNA Genen (*sRNAXcc-15/ -16/ -18*) aufwies. Dagegen hatte die Deletion der einzelnen sRNA Gene keinen Einfluss auf die *Xcc* Virulenz (1). Die entsprechenden Gene kommen im *Xcv* Stamm 85-10 nicht vor. Parallel zu den Ergebnissen dieser Arbeit wurde die Identifizierung von acht sRNAs im *Xoo* Stamm PXO99 beschrieben (118), wobei *Xoo*3, *Xoo*4 und *Xoo*6 Orthologe der *Xcv* RNAs sX14, asX4 bzw. sX1 repräsentieren (s. Kapitel 2.1.1.; (188)). Während *Xoo*4 als *trans*-kodierte sRNA (145 Nt) identifiziert wurde, handelt es sich bei asX4 in *Xcv* um eine 309 Nt-lange antisense RNA (s. Kapitel 2.1.1.; Koop. mit J. Brock; (188)). Die Deletion von einzelnen sRNA Genen in *Xoo* hatte keinen Einfluss auf die Virulenz (118). Wenngleich mögliche Ziel-mRNAs unbekannt sind, lassen Proteomanalysen vermuten, dass sRNA-*Xoo*1*,* sRNA-*Xoo*3 und sRNA-*Xoo*4 an der Regulation des Aminosäurestoffwechsels und anderen generellen zellulären Prozessen beteiligt sind (118). Für *S. maltophilia* wurden 16 sRNAs mit unbekannten

Funktionen beschrieben, von denen SmsR26 und SmsR39 Orthologe der *Xcv* sRNAs sX5 bzw. sX13 repräsentieren (179).

Neben Vertretern der *Xanthomonadaceae* wurden sRNAs in den pflanzenpathogenen Bakterien *P. syringae* pv. *tomato* Stamm DC3000, *A. tumefaciens* sowie in *Erwinia* spp. identifiziert (33,52,161,258,259). Virulenzfunktionen wurden für zwei sRNAs aus *E. amylovora* nachgewiesen (262). Zudem ist bekannt, dass RNAs der RsmB/ RsmC Familie in *E. carotovora* subsp. *carotovora* die Aktivität des translationalen Repressorproteins RsmA kontrollieren und QS, die Produktion extrazellulärer Enzyme sowie die Virulenz beeinflussen (38,39,121). RsmA ist ebenfalls für die Pathogenität von *Xcc* und *Xoo* essentiell, wenngleich die zugehörigen regulatorischen RNAs bislang unbekannt sind (34,123,264).

3.3.1. Mögliche Funktionen *cis*-kodierter asRNAs

Zwei der in dieser Arbeit identifizierten asRNAs, asX6 und PtaRNA1, werden unabhängig von HrpG und HrpX exprimiert und sind vermutlich an der Regulation der Expression von zytotoxischen Proteinen beteiligt (s. Kapitel 2.1.1. und 2.2.1.; (54,188)). Das *asX6* Gen ist im *Xcv* Plasmid pXCV183 lokalisiert und überlappt in antisense Orientierung mit der 3'-Region des *XCVd0100-XCVd0099* Operons, welches vermutlich ein Antitoxinprotein der ε-Familie bzw. ein ζ-Toxin-ähnliches Protein kodiert (s. Kapitel 2.1.1.; (188)). Die Lokalisierung der Gene lässt vermuten, dass asX6 die Expression des Toxingens posttranskriptionell beeinflussen könnte, dies wurde in dieser Arbeit jedoch nicht näher untersucht. Für das ζ-/ ε-Toxin-Antitoxin (TA) System in *Xcv* wird vermutet, dass es die Weitergabe von pXCV183 während der Zellteilung sicherstellt (215). Für *Streptococcus* wurde gezeigt, dass das ε-Antitoxin die zytotoxische Aktivität des ζ-Toxins unterdrückt. Ein Verlust des Plasmids bzw. des TA Lokus während der Zellteilung führt zum schnellen proteolytischen Abbau des Antitoxins, wohingegen das stabile Toxin eine Zelltodreaktion induziert und die Vermehrung plasmidfreier Zellen verhindert (28,119,265). Das ζ-/ ε-TA System und andere TA Systeme, in welchen ein Antitoxinprotein die Toxinaktivität hemmt, werden als Typ II-TA Systeme bezeichnet (192). Bisher ist nicht bekannt, dass asRNAs an der Regulation der Aktivität von Typ II-TA Systemen beteiligt sind.

In Typ I-TA Systemen wird die Toxinsynthese durch eine asRNA unterdrückt, welche als Antitoxin wirkt und meist konstitutiv exprimiert wird (192). Die Bindung der asRNA an die Toxin-kodierende mRNA hemmt die Ribosomenbindung und/ oder induziert den Abbau des RNA Duplex (41,56,233). Neben plasmidkodierten TA-Systemen wurden entsprechende Loci auch in den Chromosomen zahlreicher Bakterien identifiziert (57,127). In *Xcv* kodiert das *ptaRNA1* Gen („plasmid-transferred antisense RNA 1') vermutlich das RNA Antitoxin (72 Nt) eines chromosomal-lokalisierten Typ I-TA Systems (s. Kapitel 2.2.1.; (54)). Das asRNA Gen überlappt mit der RBS und der 5'-Region des *XCV2162* ORFs, welcher ein hypothetisches Protein (76 Aminosäuren; AS) mit unbekannter Funktion kodiert (s. Kapitel 2.2.1.; (54))(215). Der Lokus weist typische Merkmale eines Typ I-TA Systems

auf, da dRNA-Seq ‚reads' für PtaRNA1 im Vergleich zu *XCV2162* deutlich überrepräsentiert sind (~200 bzw. 2 ‚reads')(s. Kapitel 2.1.1. und 2.2.1.; (54,188))(66). Phylogenetische Analysen ergaben, dass der potentielle TA Lokus in zahlreichen, nur entfernt mit *Xcv* verwandten γ- und ß-Proteobakterien konserviert ist, jedoch nicht in nahe verwandten Bakterien vorkommt (s. Kapitel 2.2.1.; (54)). Zudem sind Orthologe von *ptaRNA1*/ *XCV2162* überwiegend in Nachbarschaft des *trbL* Gens kodiert, welches vermutlich am konjugalen DNA Transfer beteiligt ist. Das sporadische phylogenetische Auftreten sowie die Plasmidlokalisierung eines entsprechenden Lokus in *P. aeruginosa* lassen vermuten, dass dieses TA System seinen Ursprung in Plasmiden hat und durch horizontalen Gentransfer verbreitet wird (s. Kapitel 2.2.1.; (54)). Die meisten Typ I-Toxine, einschließlich XCV2162, sind kleine Proteine (20-65 AS) mit Transmembrandomänen (s. Kapitel 2.2.1.; (54))(57). Für die *E. coli* Proteine Hok, IbsC und ShoB wird beispielsweise angenommen, dass sie das Membranpotential der Zelle beeinträchtigen können (55,56,57,65). Einen Sonderfall stellt das *E. coli symE*/ SymR Typ I-TA System dar. *symE* kodiert vermutlich eine toxinähnliche RNA-Endonuklease und wird als Antwort auf DNA-Schädigung exprimiert (99). Bislang wurde nicht untersucht, ob *XCV2162* ein Toxin kodiert und unter welchen Bedingungen es exprimiert wird. Der Einfluss von PtaRNA1 auf die Expression von *XCV2162* könnte beispielsweise durch Mutation des *ptaRNA1* Promotors analysiert werden. Zudem könnte die ektopische Expression von *XCV2162* Hinweise auf dessen Funktion liefern.

Die HrpG-/ HrpX-abhängige Akkumulation von drei der in dieser Arbeit verifizierten asRNAs (asX1, asX4 und asX5), welche komplementär zur 3'-Region der *cis*-kodierten Gene sind (*XCV0392*, *XCV4105* bzw. *nrdB*), deutet auf mögliche Virulenzfunktionen hin (s. Kapitel 2.1.1.; (188)). Die zelluläre Funktion von XCV0392 ist nicht bekannt. XCV4105 ähnelt eukaryotischen mitochondrialen Rho GTPasen, welche u.a. am programmierten Zelltod (Apoptose) beteiligt sind (58). Interessanterweise weist *XCV4105* ein vergleichbares Expressionsmuster wie *asX4* auf und trägt vermutlich zur Virulenz von *Xcv* bei (J. Brock, C. Schmidtke und U. Bonas, unveröffentlicht). NrdB, die ß-Untereinheit der Ribonukleotid-Diphosphat Reduktase, vermittelt wahrscheinlich die Synthese von Desoxyribonukleotiden aus Ribonukleotiden und ist dadurch indirekt an der DNA Synthese beteiligt (80). Wenngleich NrdB in Bakterien konserviert ist, wurden Orthologe von *asX5* ausschließlich in *Xcv* identifiziert (s. Kapitel 2.1.1.; (188)). Dies könnte auf eine spezifische Funktion von *asX5*/ *nrdB* in der Interaktion von *Xcv* mit Wirtspflanzen hindeuten. Nachfolgende Analysen könnten die Expressionsmuster der genannten mRNAs untersuchen und regulatorische Funktionen der entsprechenden asRNAs mit Hilfe von Promotormutationen bzw. Überexpressionexperimenten ermitteln.

3.3.2. sRNAs mit potentiellen Virulenzfunktionen

Die im Rahmen dieser Arbeit identifizierten *Xcv* sRNAs sind mit Ausnahme von sX6 nicht-kodierend. Wenngleich sX6 ein Protein kodiert (80 AS)(s. Kapitel 2.1.1.; Koop. mit J. Brock; (188)), kann eine Funktion als regulatorische RNA nicht ausgeschlossen werden. Duale Funktionen wurden beispielsweise für SgrS in *E. coli* und RNAIII in *S. aureus* beschrieben (226). SgrS hemmt die Synthese des Zuckerphosphat-Transporters PtsG durch Basenpaarung mit der *ptsG* mRNA und kodiert zudem das SgrT Protein (43 AS), welches die Aktivität von Transporterproteinen inhibiert (227,237). RNAIII reguliert durch Basenpaarung die Translation von sechs mRNAs und vermittelt die Produktion von Exotoxinen. Darüber hinaus kodiert RNAIII einen sekretierten Virulenzfaktor (Hämolysin δ; 26 AS)(226).

Im Fokus dieser Arbeit stand die Identifizierung von *Xcv* sRNAs mit möglichen Virulenzfunktionen. Das HrpG- bzw. HrpX-abhängige Expressionsmuster von fünf *Xcv* sRNAs (sX4, sX5, sX8, sX11 und sX12) deutet auf eine solche Funktion hin (s. Kapitel 2.1.1.; (188)). Bislang wurde nur sX12 (78 Nt) näher untersucht. sX12 wird HrpX-abhängig exprimiert und fördert die Ausbildung wässriger Läsionen in suszeptiblen sowie die HR-Induktion in resistenten Pflanzen (s. Kapitel 2.1.1.; (188)). Die Deletion von *sX12* hat keinen Einfluss auf die Vermehrung von *Xcv in planta* oder die generelle Fähigkeit zur Typ III Sekretion *in vitro* (s. Kapitel 2.1.1.; Koop. mit J. Brock; (188)). Um die molekularen Funktionen von sX12 zu untersuchen, wurden bislang qRT-PCR Analysen von mehr als zehn bioinformatisch vorhergesagten Ziel-mRNAs sowie Proteom- und ‚Microarray' Analysen von *Xcv* Stamm 85-10 und der *sX12* Deletionsmutante durchgeführt. Diese Experimente ergaben keine signifikanten Unterschiede in den Protein- oder Transkriptmengen der getesteten Stämme (C. Schmidtke und U. Bonas, unveröffentlicht; C. Schmidtke, B. Voigt, M. Hecker und U. Bonas, unveröffentlicht; C. Schmidtke, S. Serrania, A. Becker und U. Bonas, unveröffentlicht). Bislang ist nicht bekannt, worauf die reduzierte Virulenz der *sX12* Deletionsmutante beruht. In zukünftigen Untersuchungen könnten sogenannte ‚pulse expression' Analysen durchgeführt werden. Hierbei wird die Expression der sRNA mittels eines induzierbaren Promotors kurzzeitig und stark induziert, wobei angenommen wird, dass dies vor allem direkt gebundene mRNAs beeinflusst (158,159). Durch nachfolgende RNA-Seq Analysen könnten potentielle sX12 Ziel-mRNAs identifiziert werden. Allerdings könnte sX12 auch die Translation von mRNAs modulieren, ohne die mRNA Akkumulation zu beeinflussen. Denkbar wäre auch, dass sX12 an Proteine bindet und deren Aktivität reguliert. Mögliche sX12-gebundene Proteine könnten mittels eines RNA-Epitop-markierten sRNA Derivats gereinigt werden (184).

Die Funktion von sX12 ist vermutlich mit dem T3S System verknüpft, da die Expression von sX12 mit dem T3S System ko-reguliert ist und da *sX12* Orthologe ausschließlich in *Xanthomonas* spp. vorkommen, die ein *hrp*-T3S System kodieren (s. Kapitel 2.1.1.; (188)). Hierbei könnte sX12 die Aktivität oder Expression einzelner Komponenten oder Substrate des T3S Systems beeinflussen. Eine

solche Funktion wird auch für die *Salmonella* sRNA IsrJ vermutet. IsrJ ist Teil des SPI-1 ("*Salmonella* pathogenicity island 1") Regulons und fördert durch einen bislang unbekannten Mechanismus die Translokation von SPI-1 Typ III Effektoren sowie die bakterielle Invasion von Epithelzellen (155).

Um weitere sRNAs zu identifizieren, die zur Virulenz von *Xcv* beitragen, könnte mittels "dual sequencing" das Transkriptom von Pathogen und Wirt während der Infektion analysiert werden (249,256). Denkbar wäre auch, dass bakterielle Transkripte in die Wirtszelle transloziert werden und dort die Expression pflanzlicher Abwehrgene unterdücken. Eine solche Funktion wurde erstmals kürzlich für microRNAs des pflanzenpathogenen Pilzes *Botrytis cinerea* nachgewiesen (249).

3.4. sX13 – ein neuartiger Regulator der Virulenzgenexpression

3.4.1. sX13 fördert die *hrp*-Genexpression und die Virulenz von *Xcv*

Die meisten *Xcv* Virulenzfaktoren, einschließlich *sX12*, wurden anhand ihrer Ko-Expression mit dem T3S System identifiziert (s. Kapitel 2.1.1.; (188))(24). Unerwarteterweise beeinträchtigte die Deletion des konstitutiv exprimierten *sX13* Gens die Virulenz von *Xcv* bzw. die Fähigkeit zur HR Induktion (s. Kapitel 2.3.1.; (187)). Dies beruht vermutlich nicht auf einer veränderten Fitness der Bakterien, da das *in planta* Wachstum von Mutante und Wildtypstamm vergleichbar war (s. Kapitel 2.3.1.; (187)). Expressionsanalysen ergaben, dass sX13 die Expression von *hrpX* und HrpX-kontrollierten Genen nach Anzucht von *Xcv* in *hrp*-Gen induzierendem XVM2 Medium fördert (s. Kapitel 2.3.1.; (187)). Dies lässt vermuten, dass sX13 die Aktivität des T3S Systems fördert und dadurch zur Virulenz von *Xcv* beiträgt. Interessanterweise war die mRNA Abundanz von *hrpX* sowie des HrpX-induzierten Effektorgens *xopS* auch nach Anzucht von *Xcv* in NYG Komplexmedium sX13-abhängig verändert (s. Kapitel 2.3.1.; (187)). Dies zeigt, dass HrpX, anders als bislang angenommen (189,252), bereits in geringem Maße unter nicht-*hrp*-Gen induzierenden Bedingungen aktiv ist. Dieses Ergebnis beruht vermutlich auf der höheren Sensitivität der in dieser Arbeit verwendeten "Microarray" und qRT-PCR Analysen, verglichen mit den zuvor getesteten Promotor-Reportergenfusionen (189,252).

Wie wirkt sX13 auf das *hrp*-Regulon? sX13 beeinflusste einerseits die mRNA Akkumulation von *hrpX*, hatte jedoch keinen Einfluss auf *hrpG* und die Expression translationaler *hrpG*- bzw. *hrpX*-Reportergenfusionen (s. Kapitel 2.3.1.; (187)). Dies lässt vermuten, dass sX13 indirekt auf das *hrp*-Regulon wirkt. Da die Reporterplasmide nur einen Teil der kodierenden Sequenzen von *hrpG* und *hrpX* enthielten, sollte untersucht werden, ob sX13 die Translation der Volllängen mRNAs beeinflusst. sX13 fördert vermutlich die Aktivität des HrpG Proteins, da die ektopische Expression von HrpG*, einer konstitutiv aktiven Version von HrpG (254), die Deletion von *sX13* hinsichtlich des *in planta* Phänotyps sowie der Expression von *hrpX* und HrpX-regulierter Gene kompensierte (s. Kapitel 2.3.1.; (187)). Denkbar wäre, dass sX13 die Synthese eines bislang unbekannten HrpG-aktivierenden Proteins

reguliert. HrpG gehört zur Familie der OmpR-Typ ‚response regulators', welche typischerweise durch zugehörige Sensor-Histidinkinasen phosphoryliert und aktiviert werden (255). Das *Xcv* Genom kodiert mehr als 80 solcher Proteine (215). Kürzlich wurde eine in *Xanthomonas* spp. konservierte Sensor-Histidinkinase, HpaS (‚*hrp*-associated sensor'), im *Xcc* Stamm 8004 als Regulator der HrpG Aktivität identifiziert (116). Die Mutation von *hpaS* beeinträchtigt die Phosphorylierung von HrpG sowie die Virulenz von *Xcc* (116). Zukünftige Analysen könnten untersuchen, ob in *Xcv* die Aktivität von HrpG durch HpaS kontrolliert wird bzw. inwiefern sX13 die Expression von HpaS oder anderen möglichen Sensor-Histidinkinasen beeinflusst.

Die regulatorische Rolle von sX13 ist nicht allein auf das *Xcv* T3S System beschränkt. Dies wird u.a. anhand der phylogenetischen Verbreitung von *sX13* deutlich (s. Kapitel 2.3.1.; (187)). Homologe sind in allen Vertretern der *Xanthomonadaceae* Familie konserviert, von denen einige kein *hrp*-T3S System kodieren, z.B. *Xal* und *S. maltophilia* (s. Kapitel 2.3.1.; (187)). Interessanterweise sind alle *sX13* Homologe stromabwärts des DNA Polymerase I-kodierenden *polA* Gens lokalisiert (s. Kapitel 2.3.1.; (187))(51). Dieser Lokus kodiert in *E. coli* und α-Proteobakterien die sRNA Spot42 bzw. sRNAs der αr7 Familie (45,97,173,183). sX13 weist keine Sequenzähnlichkeit zu diesen sRNAs auf, wenngleich αr7 sRNAs ‚Stem-Loops' mit ‚C'-reichen Loop-Sequenzen enthalten und sX13 strukturell ähneln (s. Kapitel 2.3.1.; (187))(45). Die Funktion von αr7 sRNAs ist nicht bekannt. Spot42 ist an der Regulation des Zuckerstoffwechsels in *E. coli* beteiligt (9,141). sX13, Spot42 und αr7 RNAs haben sich vermutlich unabhängig voneinander oder durch divergente Evolution eines gemeinsamen Vorläufers entwickelt.

3.4.2. Mögliche physiologische Funktionen von sX13

In dieser Arbeit wurden 63 potentielle sX13-Ziel mRNAs anhand veränderter Transkriptmengen in der *sX13* Deletionsmutante gegenüber dem *Xcv* Wildtypstamm identifiziert, wobei die Expression von *XCV0173*, *XCV3573* und *pilH* zwar mittels qRT-PCR, jedoch nicht durch die ‚Microarray' Analysen nachgewiesen werden konnte (s. Kapitel 2.3.1.; (187)). Dies ist auf die geringe Signalstärke der ‚Microarray' Hybridisierungsignale für diese mRNAs zurückzuführen und deutet auf eine geringe Sensitivität der ‚Microarray' Analysen hin. Folglich könnte sX13 die Akkumulation von weiteren mRNAs beeinflussen. Die ‚Microarray' Analysen erlauben keine Rückschlüsse darauf, ob sX13 direkt mit mRNAs intergagiert (s. Kapitel 3.4.5.). Da das sX13 Regulon Gene umfasst, die transkriptionelle Regulatoren kodieren, z.B. *hrpX*, *algR* und *pilH* (s. Kapitel 2.3.1.; (187)), liegt die Vermutung nahe, dass sX13 indirekt die Transkription von *Xcv* Genen beeinflusst.

Die Ergebnisse zeigen außerdem, dass sX13 die Expression des RNA-Bindeproteins Hfq hemmt (s. Kapitel 2.3.1.; (187)), welches in zahlreichen Bakterien die Stabilität und Aktivität von sRNAs kontrolliert (44). Folglich könnte sX13 durch Regulation der Hfq Synthese indirekt posttranskriptionelle Prozesse in *Xcv* beeinflussen. Die Ergebnisse von J. Brock zeigen, dass die

Mutation von *hfq* keine Auswirkung auf die Virulenz von *Xcv* oder die Abundanz von sX13 hatte, jedoch die Akkumulation der sX14 sRNA beeinträchtigte (s. Kapitel 2.3.1.; (187)). Zudem wurde nachgewiesen, dass die Mutation von *hfq* keinen Einfluss auf die sX13-abhängige Expression von mRNA*::gfp* Fusionen hat, was zeigt, dass sX13 Hfq-unabhängig agiert (s. Kapitel 2.3.1.; Koop. mit U. Abendroth und J. Brock; (187)). Bislang ist sX13 die einzige bakterielle sRNA, für die ein Einfluss auf *hfq* nachgewiesen wurde. In Übereinstimmung mit den Ergebnissen von J. Brock wurde für *Xoo* beschrieben, dass die Deletion von *hfq* die Akkumulation des sX14 Orthologs sRNA-*Xoo*3 beeinträchtigt und keinen Einfluss auf die Virulenz hat (118). Zudem wurde für *S. maltophilia* nachgewiesen, dass das sX13 Ortholog SmsR39 Hfq-unabhängig akkumuliert (179). In den meisten pathogenen Bakterien, einschließlich des Pflanzenpathogens *A. tumefaciens*, trägt Hfq zur Virulenz bei (35,257), wohingegen beispielsweise die *S. aureus* sRNA RNAIII in Hfq-unabhängiger Weise die Virulenzgenexpression kontrolliert (13).

Die in dieser Arbeit durchgeführten Expressionsanalysen lassen vermuten, dass sX13 durch Modulation der Genexpression zur Adaption von *Xcv* an sich verändernde Umweltbedingungen beiträgt (Abb. 4). Hierfür spricht, dass die sRNA unter bestimmten Wachstumsbedingungen akkumulierte und dass die Abundanz ausgewählter mRNAs, z.B. *hrpX*, mit der Abundanz von sX13 korrelierte (s. Kapitel 2.3.1.; (187)). Zudem beeinflusste sX13 in gegensätzlicher Weise und in Abhängigkeit von den Wachstumsbedingungen die Akkumulation von mRNAs, welche an der Tfp Biogenese bzw. der Flagellum-vermittelten Chemotaxis beteiligt sind (s. Kapitel 2.3.1.; (187)). Tfp vermitteln die bakterielle Fortbewegung auf festen Oberflächen, wohingegen das Flagellum eine schwimmende Fortbewegung ermöglicht (91). Für *Xanthomonas* spp. wurde beschrieben, dass Tfp und Flagellum Virulenzfunktionen erfüllen und u.a. die bakterielle Anheftung an die Blattoberfläche ermöglichen (42,128,225,240). Folglich könnte sX13 in Abhängigkeit von den Umweltbedingungen die Art der bakteriellen Fortbewegung modulieren bzw. zur Kolonisierung pflanzlicher Oberflächen durch *Xcv* beitragen.

sX13 ist vermutlich an der QS-abhängigen Regulation der Genexpression beteiligt, da die Zelldichte der *sX13* Deletionsmutante in der stationären Wachstumsphase gegenüber dem *Xcv* Wildtypstamm reduziert war (s. Kapitel 2.3.1.; (187)). In diesen Experimenten wurde sowohl die optische Dichte der Kulturen gemessen, als auch die bakterielle Zellzahl bestimmt. Die Zellzahl des *Xcv* Wildtypstamms und der *sX13* Deletionsmutante korrelierte in vergleichbarem Maße mit der optischen Dichte der Kulturen (C. Schmidtke und U. Bonas, unveröffentlicht). Ein weiterer Hinweis auf eine Funktion von sX13 in der QS-abhängigen Regulation ist die sX13-abhängige Akkumulation der *XCV2041* mRNA, welche ein Protein kodiert, das GGDEF- und EAL-Domänen aufweist (s. Kapitel 2.3.1.; (187)). Solche Proteine kontrollieren typischerweise die Menge des intrazellulären Botenmoleküls zyklisches di-GMP und modulieren dadurch die QS-abhängige Genexpression (78). Interessanterweise weist XCV2041 94% Sequenzübereinstimmung mit dem *Xcc* Protein XC2226 auf, welches als Repressor der Tfp-vermittelten Motilität beschrieben wurde (182). Studien in *Xac*, *Xcc* und *Xoc* ergaben, dass QS

die Expression des *hrp*-Regulons und die Flagellum-vermittelte Motilität beeinflusst bzw. dass HrpG die Expression von Genen des QS Systems und des Flagellarapparats moduliert (73,76,77,182,263).

Insgesamt lassen die Ergebnisse dieser Arbeit den Schluss zu, dass sX13 in *Xcv* die Aktivität verschiedener regulatorischer Netzwerke koordiniert und dadurch die Virulenzgenexpression, QS und die Motilität in Abhängigkeit von den Umweltbedingungen beeinflusst (Abb. 4)(s. Kapitel 2.3.1.; (187)). Weitere Analysen könnten untersuchen, durch welche Signalwege und Transkriptionsregulatoren die Transkription von *sX13* kontrolliert wird.

Auch in anderen pathogenen Bakterien wurden sRNAs identifiziert, welche QS und die Virulenzgenexpression kontrollieren: In *V. cholerae* akkumulieren vier homologe und redundant wirkende sRNAs (Qrr1-4) bei niedriger Zelldichte und hemmen die Translation der *hapR* mRNA (4,114). HapR kontrolliert die QS-Antwort und unterdrückt die Transkription von T3S- und anderen Virulenzgenen. Die Expression der *S. aureus* sRNA RNAIII steigt mit zunehmender Zelldichte. Durch Repression der Translation der *rot* mRNA (,repressor of toxins') hemmt RNAIII die Synthese von Außenmembranproteinen und fördert die Produktion sekretierter Toxine (15,61,151).
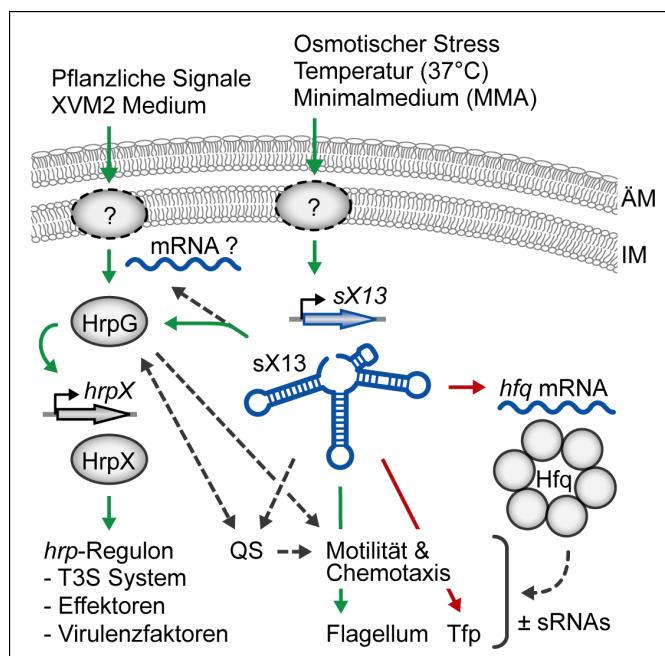
Abbildung 4. Modell physiologischer Funktionen von sX13 in *Xcv*.
Die Expression bzw. Akkumulation von sX13 wird durch extrazelluläre Stimuli und unbekannte Signalwege induziert und beeinflusst die Aktivität regulatorischer Netzwerke. sX13 fördert indirekt die Aktivität von HrpG und beeinflusst dadurch die Expression von *hrpX* sowie des *hrp*-Regulons. HrpG wird in Gegenwart pflanzlicher Signale bzw. XVM2 Medium durch unbekannte Signalwege aktiviert, wobei sX13 möglicherweise die Synthese eines HrpG-aktivierenden Proteins moduliert. sX13 hemmt vermutlich die Typ IV Pilus-vermittelte Motilität, fördert eine Flagellum-abhängige schwimmende Fortbewegung von *Xcv* und beeinflusst möglicherweise die Zelldichte-abhängige Genexpression (QS). Für andere *Xanthomonas* spp. wurde nachgewiesen, dass *hrp*- und QS-Regulon partiell überlappen und Motilitätsgene umfassen. Durch Hemmung der Hfq Synthese beeinflusst sX13 vermutlich weitere sRNA-vermittelte posttranskriptionelle Prozesse. (Umrahmte Pfeile, Wellenlinien und Kreise kennzeichnen Gene, mRNAs bzw. Proteine. Unbekannte Proteine/ Signalwege sind durch gestrichelte Kreise dargestellt. Grüne, rote und gestrichelte Pfeile kennzeichnen positive, negative bzw. vermutete Effekte. IM, innere Membran; ÄM, äußere Membran).

3.4.3. Die Aktivität von sX13 beruht auf ‚C'-reichen Loops

Die Strukturanalyse von sX13 wurde von U. Abendroth durchgeführt und ergab drei doppelsträngige Bereiche, welche einzelsträngige apikale Loops mit ‚C'-reichen (‚4C'/ ‚5C') Motiven enthalten (s. Kapitel 2.3.1.; (187)). Wenngleich die Sequenzen auf redundante Funktionen der sX13 Loops hindeuten, ergaben die von U. Abendroth durchgeführten Komplementationsexperimente mit plasmidkodierten *sX13* Derivaten, dass die Loops 2 und 3, jedoch nicht Loop 1 oder die unstrukturierte 5'-Region, zur Virulenz von *Xcv* beitragen (s. Kapitel 2.3.1.; (187)). qRT-PCR Analysen und Experimente mit translationalen mRNA*::gfp* Fusionen ergaben, dass sX13 mittels Loop 2 die mRNA Akkumulation von *XCV3927* und *hfq* sowie die Synthese von XCV3927::GFP und Hfq::GFP hemmt (s. Kapitel 2.3.1.; (187)). Dagegen hemmen vermutlich mehrere sX13 Loops die mRNA Akkumulation von *pilH* bzw. die Synthese von PilH::GFP (s. Kapitel 2.3.1.; (187)). Die Expression des plasmidkodierten *sX13* Gens sowie der *sX13* Loopmutanten wurde mittels Northern Blot Analysen nachgewiesen, wobei im Gegensatz zum chromosomal-lokalisierten *sX13* Gen mehrere Hybridisierungssignale detektiert wurden (s. Kapitel 2.3.1.; (187)). Diese Signale waren jedoch identisch für die verschiedenen sX13 Derivate und beruhen vermutlich auf einer ineffizienten Termination der Transkription der plasmidkodierten Gene.

sX13 beeinflusste die Synthese der genannten GFP-Fusionsproteine in ähnlichem Maße wie die mRNA Abundanz der entsprechenden chromosomal-kodierten Gene (s. Kapitel 2.3.1.; (187)). Diese Ergebnisse ließen zunächst vermuten, dass sX13 die Akkumulation potentieller Ziel-mRNAs, jedoch nicht deren Translation beeinflusst. Dagegen spricht, dass die Fluoreszenz des XCV3927::GFP Proteins, jedoch nicht die Abundanz der *XCV3927::gfp* mRNA sX13-abhängig verändert war (s. Kapitel 2.3.1.; (187)). Folglich beeinflusst sX13 die mRNA Akkumulation und Translation von *XCV3927* unabhängig voneinander. Die Ergebnisse deuten darauf hin, dass weitere Sequenzen des *XCV3927* ORFs, welche nicht im *XCV3927::gfp* Derivat enthalten waren, für die sX13-abhängige mRNA Akkumulation essentiell sind. Dies könnte durch Expressionsanalysen von *gfp*-Fusionen der Volllängen-mRNAs untersucht werden. Aufgrund der zuvor erwähnten Rifampicin Resistenz der verwendeten *Xcv* Stämme steht der Nachweis aus, dass sX13 die Stabilität, d.h. die Halbwertszeit, möglicher Ziel-mRNAs beeinflusst. Da sRNAs häufig den Abbau von Ziel-mRNAs durch RNase E oder RNase III induzieren (s. Kapitel 1.3.4.1.), könnte der Einfluss von sX13 auf die Stabilität von Ziel-mRNAs durch Mutation von *sX13* sowie der entsprechenden RNase-kodierenden Gene untersucht werden. Ähnlich zu den Ergebnissen für sX13/ *XCV3927* in *Xcv* wurde für *Salmonella* beschrieben, dass die sRNA RyhB die RNase E-vermittelte Spaltung der *sodB* mRNA in rund 350-Nt Entfernung stromabwärts von der sRNA-Bindestelle induziert (169). Zudem wurde für RyhB in *E. coli* nachgewiesen, dass die Repression der *sodB* Translation unabhängig von der RNaseE-vermittelten mRNA Degradation erfolgt (145).

3.4.4. ‚G'-reiche mRNA Motive vermitteln die sX13-abhängige Genexpression

Wie wirkt sX13 auf mögliche Ziel-mRNAs? sX13 hemmte die Synthese von XCV3927::GFP und PilH::GFP Fusionsproteinen, welche unter Kontrolle des sX13-unabhängigen *lac*-Promotors exprimiert wurden (s. Kapitel 2.3.1.; (187)). Dies zeigt, dass sX13 als posttranskriptioneller Repressor wirkt. Da sRNAs auch die Translation von mRNAs fördern können (60), sollte untersucht werden, ob sX13 die Expression positiv regulierter mRNAs posttranskriptionell beeinflusst.

Die bioinformatische Analyse der ‚Microarray' Daten ergab, dass 70% der durch sX13 negativ beeinflussten mRNAs, einschließlich *XCV3927*, *hfq* und *pilH*, ein oder mehrere ‚GGGG' (‚4G')-Motive in den 5'-Regionen aufweisen (s. Kapitel 2.3.1.; (187)). Dies lässt vermuten, dass sX13 über ‚C'-reiche Loops mit ‚G'-reichen Motiven in potentiellen Ziel-mRNAs interagiert. sX13 weist strukturelle Ähnlichkeit zu RNAIII und anderen sRNAs aus *S. aureus* auf, welche in Loops lokalisierte ‚UCCC'-Motive enthalten und mit ‚G'-reichen Sequenzen in den 5'-UTRs von Ziel-mRNAs interagieren (s. Kapitel 2.3.1.; Koop. mit U. Abendroth; (187))(62). Zudem ist die Loop 3 Sequenz von sX13 (‚UCCCCCU') identisch zu einem Teil der Loop-Sequenz der *H. pylori* sRNA HPnc5490 (s. Kapitel 2.3.1.; (187)). Diese hemmt wahrscheinlich durch komplementäre Basenpaarung die Synthese des Chemotaxisregulators TlpB (196).

Die meisten bakteriellen sRNAs hemmen die Initiation der Translation von Ziel-mRNAs (245), wobei Mutationen in sRNA oder mRNA zum Verlust der translationalen Repression führen (19,224). Eine direkte Interaktion wird meist dann angenommen, wenn die sRNA-vermittelte Repression durch komplementäre Sequenzaustausche in sRNA und mRNA wiederhergestellt wird (19,210,224). Eine Rolle in der Hemmung der Translationsinitiation wurde zunächst auch für sX13 angenommen (Abb. 5A), da die Mutation der sX13 Loops mit einer verstärkten Synthese der entsprechenden GFP-Fusionsproteine verbunden war (s. Kapitel 2.3.1.; (187)). Ein solches Modell wird jedoch nicht durch die übrigen Ergebnisse gestützt. Entgegen der Erwartung führte die Mutation der ‚4G'-mRNA Motive nicht zu einer verstärkten Synthese von XCV3927::GFP und Hfq::GFP. Stattdessen war die Expression der mutierten mRNA*::gfp* Fusionen in An- und Abwesenheit von sX13 oder Derivaten ähnlich reduziert, wie die Expression der nicht-mutierten mRNA*::gfp* Fusionen in Gegenwart von sX13 (s. Kapitel 2.3.1.; (187)). Zudem führte die Mutation des ‚4G'-Motivs in der *XCV3927::gfp* mRNA, welche unter Kontrolle des *lac*-Promotors exprimiert wurde, zu einer reduzierten und sX13-unabhängigen Expression des Fusionsproteins (s. Kapitel 2.3.1.; (187)). Diese Ergebnisse lassen folgende Schlussfolgerungen zu: (i) ‚C'-reiche sX13 Loops und ‚G'-reiche Motive in möglichen Ziel-mRNAs sind für die Aktivität von sX13 essentiell. (ii) sX13 beeinflusst vermutlich nicht die Translationsinitiation von *XCV3927::gfp* und *hfq::gfp*. (iii) ‚4G'-Motive in möglichen sX13-Ziel mRNAs können die Translation fördern, d.h. wirken als translationale Verstärker (‚enhancer'). Wenngleich die in sX13 und mRNAs eingeführten Mutationen die Sequenzkomplementarität

wiederherstellen, erlauben die Ergebnisse keine Rückschlüsse darauf, ob sX13 direkt mit mRNAs interagiert (s. Kapitel 3.4.5.).

Insgesamt können die erhaltenen Ergebnisse durch ein Modell erklärt werden, in welchem sX13 an ‚G'-reiche Motive in mRNAs bindet und dadurch die Bindung eines Translations-fördernden Faktors verhindert (Abb. 5B). Ein solcher Faktor könnte eine sRNA, ein Protein oder das Ribosom selbst sein. Die *Xcv* sRNA sX5 ähnelt sX13 und weist zwei in einem einzelnen Loop lokalisierte ‚4C'-Motive auf (s. Kapitel 2.1.1.; (188)). Daher sollte untersucht werden, inwiefern sich mögliche Ziel-mRNAs von sX13 und sX5 überschneiden bzw. ob die Mutagenese oder Überexpression von sX5 die sX13 Aktivität beeinflusst. Alles in allem ist eine Rolle des Ribosoms in der Bindung ‚G'-reicher mRNA Motive naheliegend, da 5% aller proteinkodierenden *Xcv* Gene, einschließlich *pilH*, ein ‚4G'- oder ‚5G'-Motiv an den Positionen 8 bis 15 stromaufwärts des TLS aufweisen (s. Kapitel 2.3.1.; (187)). Diese Hypothese wird dadurch gestützt, dass die Mutation des ‚5G'-Motivs zum Verlust der PilH::GFP Synthese führte (s. Kapitel 2.3.1.; (187)). Wenngleich *Xcv* mRNAs keine konservierte SD-Sequenz aufweisen (s. Kapitel 3.1.1.), könnte die Anti-SD Sequenz der 16S rRNA theoretisch durch ‚U-G' (‚non-Watson-Crick') Basenpaarungen mit ‚4G'- oder ‚5G'-mRNA Motiven die Ribosomenbindung vermitteln.

Insgesamt zeigen die Ergebnisse, dass ‚G'-reiche mRNA Motive eine sX13-abhängige Expression vermitteln und abhängig von ihrer Lokalisierung als translationale ‚enhancer' wirken können (s. Kapitel 2.3.1.; (187)). Die Gegenwart eines ‚G'-reichen Motivs ist offenbar nicht allein ausschlaggebend für eine sX13-abhängige mRNA Expression, da *XCV0612* zwar ähnlich wie *pilH* ein ‚5G'-Motiv an der RBS Position trägt, jedoch sX13-unabhängig exprimiert wird (s. Kapitel 2.3.1.; (187)). Denkbar wäre, dass neben ‚G'-reichen Motiven weitere regulatorische mRNA Sequenzen für eine sX13-abhängige Expression erforderlich sind. Ebenso könnten mRNA Sekundärstrukturen zur Bindung von sX13 beitragen. Beispielsweise bindet RNAIII in *S. aureus* mittels zwei ‚C'-reicher Loops an zwei in Loops lokalisierte ‚G'-reiche Motive der *rot* mRNA (15,61). Das für sX13 vorgeschlagene Modell (Abb. 5B) ähnelt der Funktionsweise der *Salmonella* sRNA GcvB, welche die Translation von Ziel-mRNAs durch Bindung an ‚CA'-reiche ‚enhancer' Motive hemmt (195,197). In *Salmonella* und *E. coli* fördern ‚CA'-reiche mRNA Sequenzen die Ribosomenbindung und Translation, unabhängig von ihrer Lokalisierung stromaufwärts oder stromabwärts des TLS (132,195).
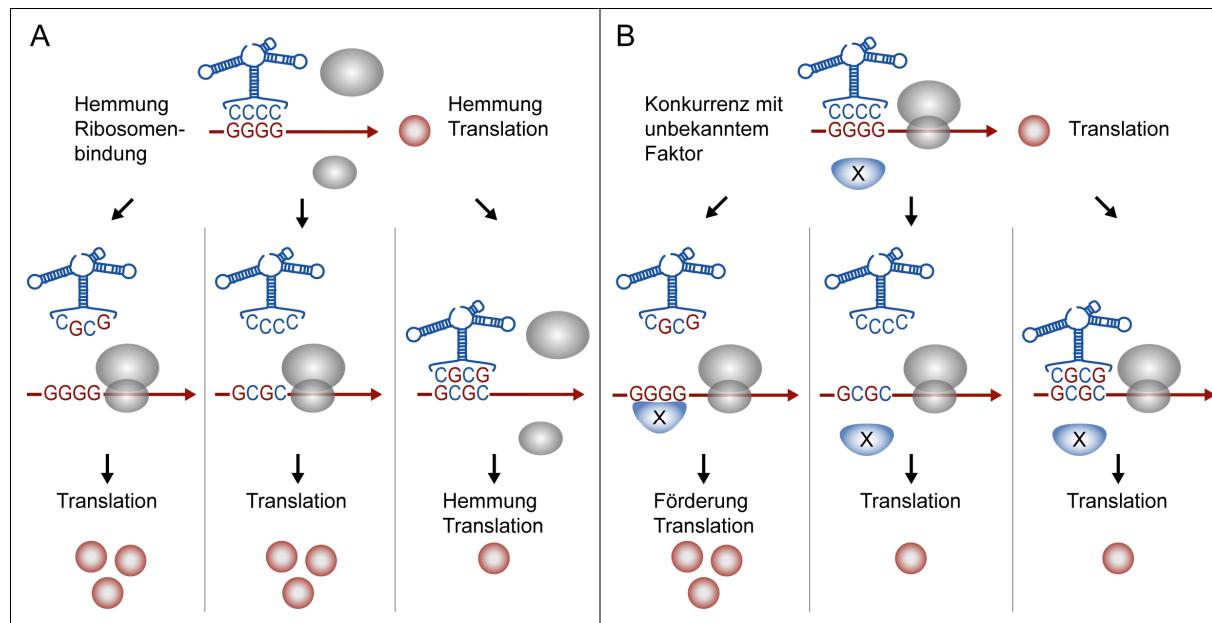
Abbildung 5. Mögliche Modelle der sX13-vermittelten Repression der mRNA Expression.
(A) Inhibierung der Translationsinitiation. Die Interaktion ‚C'-reicher sX13 Loops (blau) mit ‚G'-reichen Motiven in den 5'-Regionen von mRNAs (rot) hemmt die Bindung des Ribosoms (grau) und die Synthese des Proteins (roter Kreis)(oben). Die translationale Repression wird durch Mutation der sX13 Loops bzw. der ‚G'-reichen mRNA Motive aufgehoben (linke Seite und Mitte) und durch komplementäre Sequenzaustausche in sX13 und Ziel-mRNAs wiederhergestellt (rechte Seite). (B) sX13 hemmt die Aktivität ‚G'-reicher ‚enhancer' Elemente. Die Interaktion von sX13 Loops mit Ziel-mRNAs hemmt die Bindung eines unbekannten translationsfördernden Faktors (X) an ‚G'-reiche mRNA Motive (oben). Die Mutation der sX13 Loops führt zum Verlust der sX13-mRNA Interaktion, wodurch die Bindung des unbekannten Faktors ermöglicht und die mRNA Translation gefördert wird (linke Seite). Die Mutation der ‚G'-reichen Motive beeinträchtigt die Translation der mRNA (Mitte) und kann nicht durch komplementäre Sequenzaustausche in sX13 kompensiert werden (rechte Seite).

3.4.5. Mögliche weiterführende Untersuchungen an sX13

Die Ergebnisse dieser Arbeit werfen die Frage auf, wie die ‚C'-reichen Loops von sX13, insbesondere Loop 3, zur beobachteten Virulenzfunktion beitragen (s. Kapitel 2.3.1.; U. Abendroth; (187)). Diesbezüglich könnte untersucht werden, ob die Mutation der sX13 Loops die *hrp*-Gen Expression während der *Xcv* Infektion der Wirtspflanze beeinflusst. Durch den wechselseitigen Austausch der sX13 Loop-Sequenzen, Komplementationsanalysen mit einzelnen ‚Stem-Loop' Strukturen oder durch Mutation weiterer Nukleotide der Loops könnte untersucht werden, ob die Funktionalität der sX13 ‚Stem-Loops' durch ihre Position in der sRNA bzw. durch die Sequenzumgebung der ‚4C'-Motive bedingt wird. Zudem könnten ‚pulse expression' Analysen mit sX13 und Loopmutanten durchgeführt werden, um potentiell direkt-gebundene und Loop-spezifische Ziel-mRNAs zu identifizieren (s. Kapitel 3.3.2.).

Die Kenntniss von direkt gebundenen mRNAs ist essentiell für das Verständnis der Funktionsweise von sX13. Durch RNase-Schutzexperimente und sogenannte ‚gel-shift' bzw. ‚toeprint' Analysen (195,197) könnte untersucht werden, ob sX13 *in vitro* mit mRNAs interagiert bzw. ob sX13 die

Bindung von 30S Ribosomenuntereinheiten an mRNAs beeinflusst. Da acht der potentiellen sX13-Ziel mRNAs mindestens zwei ‚4G'-Motive nahe des TLS aufweisen (s. Kapitel 2.3.1.; (187)), könnte untersucht werden, ob sX13 über mehrere Loops mit multiplen ‚4G'-Motiven in bestimmten mRNAs interagiert. Eine solche Funktionsweise wurde für die sRNAs RNAIII in *S. aureus* und OxyS in *E. coli* nachgewiesen und trägt vermutlich zur Spezifität der sRNA-mRNA Interaktion bei (2,15,61). Mittels einer RNA-Epitop-markierten sRNA könnten potentielle RNA- und Proteininteraktoren von sX13 identifiziert werden (184). Dies ist insbesondere relevant, da sX13 Hfq-unabhängig agiert (s. Kapitel 2.3.1.; Koop. mit J. Brock und U. Abendroth; (187)). Folglich könnten bislang unbekannte RNA-Bindeproteine Hfq-ähnliche Funktionen in *Xcv* erfüllen. Ein solcher Fall wurde kürzlich für *S. meliloti* beschrieben (157).

4. Literaturverzeichnis

1. **An, S.Q., Febrer, M., McCarthy, Y., Tang, D.J., Clissold, L., Kaithakottil, G., Swarbreck, D., Tang, J.L., Rogers, J., Dow, J.M. and Ryan, R.P.** (2013) High-resolution transcriptional analysis of the regulatory influence of cell-to-cell signalling reveals novel genes that contribute to *Xanthomonas* phytopathogenesis. *Mol. Microbiol.*, **88**, 1058-1069.

2. **Argaman, L. and Altuvia, S.** (2000) *fhlA* repression by OxyS RNA: kissing complex formation at two sites results in a stable antisense-target RNA complex. *J. Mol. Biol.*, **300**, 1101-1112.

3. **Babitzke, P. and Romeo, T.** (2007) CsrB sRNA family: sequestration of RNA-binding regulatory proteins. *Curr. Opin. Microbiol.*, **10**, 156-163.

4. **Bardill, J.P., Zhao, X. and Hammer, B.K.** (2011) The *Vibrio cholerae* quorum sensing response is mediated by Hfq-dependent sRNA/mRNA base pairing interactions. *Mol. Microbiol.*, **80**, 1381-1394.

5. **Barrick, J.E. and Breaker, R.R.** (2007) The distributions, mechanisms, and structures of metabolite-binding riboswitches. *Genome Biol.*, **8**, R239.

6. **Barrick, J.E., Corbino, K.A., Winkler, W.C., Nahvi, A., Mandal, M., Collins, J., Lee, M., Roth, A., Sudarsan, N., Jona, I., Wickiser, J.K. and Breaker, R.R.** (2004) New RNA motifs suggest an expanded scope for riboswitches in bacterial genetic control. *Proc. Natl. Acad. Sci. U.S.A.*, **101**, 6421-6426.

7. **Beaume, M., Hernandez, D., Farinelli, L., Deluen, C., Linder, P., Gaspin, C., Romby, P., Schrenzel, J. and Francois, P.** (2010) Cartography of methicillin-resistant *S. aureus* transcripts: detection, orientation and temporal expression during growth phase and stress conditions. *PLoS ONE*, **5**, e10725.

8. **Becker, A. and Vorhölter, F.-J.** (2009) Xanthan biosynthesis by *Xanthomonas* bacteria: an overview of the current biochemical and genomic data. Caister Academic Press, Norfolk, UK.

9. **Beisel, C.L. and Storz, G.** (2011) The base-pairing RNA Spot 42 participates in a multioutput feedforward loop to help enact catabolite repression in *Escherichia coli. Mol. Cell*, **41**, 286-297.

10. **Bentley, D.R.Balasubramanian, S.Swerdlow, H.P.Smith, G.P.Milton, J.Brown, C.G.Hall, K.P.Evers, D.J.Barnes, C.L.Bignell, H.R.Boutell, J.M.Bryant, J.Carter, R.J.Keira Cheetham, R.Cox, A.J.Ellis, D.J.Flatbush, M.R.Gormley, N.A.Humphray, S.J.Irving, L.J.Karbelashvili, M.S.Kirk, S.M.Li, H.Liu, X.Maisinger, K.S.Murray, L.J.Obradovic, B.Ost, T.Parkinson, M.L.Pratt, M.R.Rasolonjatovo, I.M.Reed, M.T.Rigatti, R.Rodighiero, C.Ross, M.T.Sabot, A.Sankar, S.V.Scally, A.Schroth, G.P.Smith, M.E.Smith, V.P.Spiridou, A.Torrance, P.E.Tzonev, S.S.Vermaas, E.H.Walter, K.Wu, X.Zhang, L.Alam, M.D.Anastasi, C.Aniebo, I.C.Bailey, D.M.Bancarz, I.R.Banerjee, S.Barbour, S.G.Baybayan, P.A.Benoit, V.A.Benson, K.F.Bevis, C.Black, P.J.Boodhun, A.Brennan, J.S.Bridgham, J.A.Brown, R.C.Brown, A.A.Buermann, D.H.Bundu, A.A.Burrows, J.C.Carter, N.P.Castillo, N.Chiara, E.C.M.Chang, S.Neil Cooley, R.Crake, N.R.Dada, O.O.Diakoumakos, K.D.Dominguez-Fernandez, B.Earnshaw, D.J.Egbujor, U.C.Elmore, D.W.Etchin, S.S.Ewan, M.R.Fedurco, M.Fraser, L.J.Fuentes Fajardo, K.V.Scott Furey, W.George, D.Gietzen, K.J.Goddard, C.P.Golda, G.S.Granieri, P.A.Green, D.E.Gustafson, D.L.Hansen, N.F.Harnish, K.Haudenschild, C.D.Heyer, N.I.Hims, M.M.Ho, J.T.Horgan, A.M.Hoschler, K.Hurwitz, S.Ivanov, D.V.Johnson, M.Q.James, T.Huw Jones, T.A.Kang, G.D.Kerelska, T.H.Kersey, A.D.Khrebtukova, I.Kindwall, A.P.Kingsbury, Z.Kokko-Gonzales, P.I.Kumar, A.Laurent, M.A.Lawley, C.T.Lee, S.E.Lee, X.Liao, A.K.Loch, J.A.Lok, M.Luo, S.Mammen, R.M.Martin, J.W.McCauley, P.G.McNitt, P.Mehta, P.Moon, K.W.Mullens, J.W.Newington, T.Ning, Z.Ling Ng, B.Novo, S.M.O'Neill, M.J.Osborne, M.A.Osnowski, A.Ostadan, O.Paraschos, L.L.Pickering, L.Pike, A.C.Chris Pinkard, D.Pliskin, D.P.Podhasky, J.Quijano, V.J.Raczy, C.Rae, V.H.Rawlings, S.R.Chiva Rodriguez, A.Roe, P.M.Rogers, J.Rogert Bacigalupo, M.C.Romanov, N.Romieu, A.Roth, R.K.Rourke, N.J.Ruediger, S.T.Rusman, E.Sanches-Kuiper, R.M.Schenker, M.R.Seoane, J.M.Shaw, R.J.Shiver, M.K.Short, S.W.Sizto, N.L.Sluis, J.P.Smith, M.A.Ernest Sohna Sohna, J.Spence, E.J.Stevens, K.Sutton, N.Szajkowski, L.Tregidgo, C.L.Turcatti, G.Vandevondele, S.Verhovsky, Y.Virk, S.M.Wakelin, S.Walcott, G.C.Wang, J.Worsley, G.J.Yan, J.Yau, L.Zuerlein, M.Mullikin, J.C.Hurles, M.E.McCooke, N.J.West, J.S.Oaks, F.L.Lundberg, P.L.Klenerman, D.Durbin, R. and Smith, A.J.** (2008) Accurate whole human genome sequencing using reversible terminator chemistry. *Nature*, **456**, 53-59.

11. **Birch, R.G.** (2001) *Xanthomonas albilineans* and the antipathogenesis approach to disease control. *Mol. Plant Pathol.*, **2**, 1-11.

12. **Böhme, K., Steinmann, R., Kortmann, J., Seekircher, S., Heroven, A.K., Berger, E., Pisano, F., Thiermann, T., Wolf-Watz, H., Narberhaus, F. and Dersch, P.** (2012) Concerted actions of a

thermo-labile regulator and a unique intergenic RNA thermosensor control *Yersinia* virulence. *PLoS Pathog.*, **8**, e1002518.

13. **Bohn, C., Rigoulay, C. and Bouloc, P.** (2007) No detectable effect of RNA-binding protein Hfq absence in *Staphylococcus aureus*. *BMC Microbiol.*, **7**, 1-9.

14. **Bohn, C., Rigoulay, C., Chabelskaya, S., Sharma, C.M., Marchais, A., Skorski, P., Borezee-Durant, E., Barbet, R., Jacquet, E., Jacq, A., Gautheret, D., Felden, B., Vogel, J. and Bouloc, P.** (2010) Experimental discovery of small RNAs in *Staphylococcus aureus* reveals a riboregulator of central metabolism. *Nucleic Acids Res.*, **38**, 6620-6636.

15. **Boisset, S., Geissmann, T., Huntzinger, E., Fechter, P., Bendridi, N., Possedko, M., Chevalier, C., Helfer, A.C., Benito, Y., Jacquier, A., Gaspin, C., Vandenesch, F. and Romby, P.** (2007) *Staphylococcus aureus* RNAIII coordinately represses the synthesis of virulence factors and the transcription regulator Rot by an antisense mechanism. *Genes Dev.*, **21**, 1353-1366.

16. **Bonas, U., Schulte, R., Fenselau, S., Minsavage, G.V., Staskawicz, B.J. and Stall, R.E.** (1991) Isolation of a gene-cluster from *Xanthomonas campestris* pv. *vesicatoria* that determines pathogenicity and the hypersensitive response on pepper and tomato. *Mol. Plant Microbe Interact.*, **4**, 81-88.

17. **Boni, I.V., Isaeva, D.M., Musychenko, M.L. and Tzareva, N.V.** (1991) Ribosome-messenger recognition: mRNA target sites for ribosomal protein S1. *Nucleic Acids Res.*, **19**, 155-162.

18. **Bosl, M. and Kersten, H.** (1991) A novel RNA product of the *tyrT* operon of *Escherichia coli*. *Nucleic Acids Res.*, **19**, 5863-5870.

19. **Bouvier, M., Sharma, C.M., Mika, F., Nierhaus, K.H. and Vogel, J.** (2008) Small RNA binding to 5' mRNA coding region inhibits translational initiation. *Mol. Cell*, **32**, 827-837.

20. **Brantl, S.** (2007) Regulatory mechanisms employed by *cis*-encoded antisense RNAs. *Curr. Opin. Microbiol.*, **10**, 102-109.

21. **Brock, J.E., Pourshahian, S., Giliberti, J., Limbach, P.A. and Janssen, G.R.** (2008) Ribosomes bind leaderless mRNA in *Escherichia coli* through recognition of their 5'-terminal AUG. *RNA*, **14**, 2159-2169.

22. **Büttner, D.** (2012) Protein export according to schedule: architecture, assembly, and regulation of type III secretion systems from plant- and animal-pathogenic bacteria. *Microbiol. Mol. Biol. Rev.*, **76**, 262-310.

23. **Büttner, D. and Bonas, U.** (2002) Getting across--bacterial type III effector proteins on their way to the plant cell. *EMBO J.*, **21**, 5313-5322.

24. **Büttner, D. and Bonas, U.** (2010) Regulation and secretion of *Xanthomonas* virulence factors. *FEMS Microbiol. Rev.*, **34**, 107-133.

25. **Büttner, D. and He, S.Y.** (2009) Type III protein secretion in plant pathogenic bacteria. *Plant Physiol.*, **150**, 1656-1664.

26. **Büttner, D., Nennstiel, D., Klüsener, B. and Bonas, U.** (2002) Functional analysis of HrpF, a putative type III translocon protein from *Xanthomonas campestris* pv. vesicatoria. *J. Bacteriol.*, **184**, 2389-2398.

27. **Büttner, D., Noël, L., Stuttmann, J. and Bonas, U.** (2007) Characterization of the nonconserved *hpaB-hrpF* region in the *hrp* pathogenicity island from *Xanthomonas campestris* pv. *vesicatoria*. *Mol. Plant Microbe Interact.*, **20**, 1063-1074.

28. **Camacho, A.G., Misselwitz, R., Behlke, J., Ayora, S., Welfle, K., Meinhart, A., Lara, B., Saenger, W., Welfle, H. and Alonso, J.C.** (2002) *In vitro* and *in vivo* stability of the $\varepsilon_2\zeta_2$ protein complex of the broad host-range *Streptococcus pyogenes* pSM19035 addiction system. *Biol. Chem.*, **383**, 1701-1713.

29. **Canonne, J., Marino, D., Jauneau, A., Pouzet, C., Briere, C., Roby, D. and Rivas, S.** (2011) The *Xanthomonas* type III effector XopD targets the *Arabidopsis* transcription factor MYB30 to suppress plant defense. *Plant Cell*, **23**, 3498-3511.

30. **Canonne, J., Marino, D., Noël, L.D., Arechaga, I., Pichereaux, C., Rossignol, M., Roby, D. and Rivas, S.** (2010) Detection and functional characterization of a 215 amino acid N-terminal extension in the *xanthomonas* type III effector XopD. *PLoS ONE*, **5**, e15773.

31. **Caron, M.P., Lafontaine, D.A. and Massé, E.** (2010) Small RNA-mediated regulation at the level of transcript stability. *RNA Biol*, **7**, 140-144.

32. **Carpousis, A.J., Luisi, B.F. and McDowall, K.J.** (2009) Endonucleolytic initiation of mRNA decay in *Escherichia coli*. *Prog Mol Biol Transl Sci*, **85**, 91-135.

33. **Chai, Y. and Winans, S.C.** (2005) A small antisense RNA downregulates expression of an essential replicase protein of an *Agrobacterium tumefaciens* Ti plasmid. *Mol. Microbiol.*, **56**, 1574-1585.

34. **Chao, N.X., Wei, K., Chen, Q., Meng, Q.L., Tang, D.J., He, Y.Q., Lu, G.T., Jiang, B.L., Liang, X.X., Feng, J.X., Chen, B. and Tang, J.L.** (2008) The *rsmA*-like gene *rsmA_{Xcc}* of *Xanthomonas campestris* pv. *campestris* is involved in the control of various cellular processes, including pathogenesis. *Mol. Plant Microbe Interact.*, **21**, 411-423.

35. **Chao, Y. and Vogel, J.** (2010) The role of Hfq in bacterial pathogens. *Curr. Opin. Microbiol.*, **13**, 24-33.

36. **Chen, X.L., Tang, D.J., Jiang, R.P., He, Y.Q., Jiang, B.L., Lu, G.T. and Tang, J.L.** (2011) sRNA-*Xcc*1, an integron-encoded transposon- and plasmid-transferred *trans*-acting sRNA, is under the positive control of the key virulence regulators HrpG and HrpX of *Xanthomonas campestris* pathovar *campestris*. *RNA Biol*, **8**, 947-953.

37. **Crossman, L.C., Gould, V.C., Dow, J.M., Vernikos, G.S., Okazaki, A., Sebaihia, M., Saunders, D., Arrowsmith, C., Carver, T., Peters, N., Adlem, E., Kerhornou, A., Lord, A., Murphy, L., Seeger, K., Squares, R., Rutter, S., Quail, M.A., Rajandream, M.A., Harris, D., Churcher, C., Bentley, S.D., Parkhill, J., Thomson, N.R. and Avison, M.B.** (2008) The complete genome, comparative and functional analysis of *Stenotrophomonas maltophilia* reveals an organism heavily shielded by drug resistance determinants. *Genome Biol.*, **9**, 1-13.

38. **Cui, Y., Chatterjee, A., Liu, Y., Dumenyo, C.K. and Chatterjee, A.K.** (1995) Identification of a global repressor gene, *rsmA*, of *Erwinia carotovora* subsp. carotovora that controls extracellular enzymes, N-(3-oxohexanoyl)-L-homoserine lactone, and pathogenicity in soft-rotting *Erwinia* spp. *J. Bacteriol.*, **177**, 5108-5115.

39. **Cui, Y., Chatterjee, A., Yang, H. and Chatterjee, A.K.** (2008) Regulatory network controlling extracellular proteins in *Erwinia carotovora* subsp. carotovora: FlhDC, the master regulator of flagellar genes, activates *rsmB* regulatory RNA production by affecting *gacA* and *hexA* (*lrhA*) expression. *J. Bacteriol.*, **190**, 4610-4623.

40. **da Silva, A.C., Ferro, J.A., Reinach, F.C., Farah, C.S., Furlan, L.R., Quaggio, R.B., Monteiro-Vitorello, C.B., Van Sluys, M.A., Almeida, N.F., Alves, L.M., do Amaral, A.M., Bertolini, M.C., Camargo, L.E., Camarotte, G., Cannavan, F., Cardozo, J., Chambergo, F., Ciapina, L.P., Cicarelli, R.M., Coutinho, L.L., Cursino-Santos, J.R., El-Dorry, H., Faria, J.B., Ferreira, A.J., Ferreira, R.C., Ferro, M.I., Formighieri, E.F., Franco, M.C., Greggio, C.C., Gruber, A., Katsuyama, A.M., Kishi, L.T., Leite, R.P., Lemos, E.G., Lemos, M.V., Locali, E.C., Machado, M.A., Madeira, A.M., Martinez-Rossi, N.M., Martins, E.C., Meidanis, J., Menck, C.F., Miyaki, C.Y., Moon, D.H., Moreira, L.M., Novo, M.T., Okura, V.K., Oliveira, M.C., Oliveira, V.R., Pereira, H.A., Rossi, A., Sena, J.A., Silva, C., de Souza, R.F., Spinola, L.A., Takita, M.A., Tamura, R.E., Teixeira, E.C., Tezza, R.I., Trindade dos Santos, M., Truffi, D., Tsai, S.M., White, F.F., Setubal, J.C. and Kitajima, J.P.** (2002) Comparison of the genomes of two *Xanthomonas* pathogens with differing host specificities. *Nature*, **417**, 459-463.

41. **Darfeuille, F., Unoson, C., Vogel, J. and Wagner, E.G.** (2007) An antisense RNA inhibits translation by competing with standby ribosomes. *Mol. Cell*, **26**, 381-392.

42. **Das, A., Rangaraj, N. and Sonti, R.V.** (2009) Multiple adhesin-like functions of *Xanthomonas oryzae* pv. *oryzae* are involved in promoting leaf attachment, entry, and virulence on rice. *Mol. Plant Microbe Interact.*, **22**, 73-85.

43. **Davis, B.M. and Waldor, M.K.** (2007) RNase E-dependent processing stabilizes MicX, a *Vibrio cholerae* sRNA. *Mol. Microbiol.*, **65**, 373-385.

44. **De Lay, N., Schu, D.J. and Gottesman, S.** (2013) Bacterial small RNA-based negative regulation: Hfq and its accomplices. *J. Biol. Chem.*, **288**, 7996-8003.

45. **del Val, C., Romero-Zaliz, R., Torres-Quesada, O., Peregrina, A., Toro, N. and Jimenez-Zurdo, J.I.** (2011) A survey of sRNA families in a-proteobacteria. *RNA Biol*, **9**, 119-129.

46. **Doidge, E.M.** (1921) A tomato canker. *Ann. Appl. Biol.*, **7**, 407-430.

47. **Dow, J.M., Crossman, L., Findlay, K., He, Y.Q., Feng, J.X. and Tang, J.L.** (2003) Biofilm dispersal in *Xanthomonas campestris* is controlled by cell-cell signaling and is required for full virulence to plants. *Proc. Natl. Acad. Sci. U.S.A.*, **100**, 10995-11000.

48. **Dow, J.M., Fouhy, Y., Lucey, J.F. and Ryan, R.P.** (2006) The HD-GYP domain, cyclic di-GMP signaling, and bacterial virulence to plants. *Mol. Plant Microbe Interact.*, **19**, 1378-1384.

49. **Ellis, J.C. and Brown, J.W.** (2009) The RNase P family. *RNA Biol*, **6**, 362-369.

50. **Epshtein, V., Mironov, A.S. and Nudler, E.** (2003) The riboswitch-mediated control of sulfur metabolism in bacteria. *Proc. Natl. Acad. Sci. U.S.A.*, **100**, 5052-5056.

51. **Fijalkowska, I.J., Schaaper, R.M. and Jonczyk, P.** (2012) DNA replication fidelity in *Escherichia coli*: a multi-DNA polymerase affair. *FEMS Microbiol. Rev.*, **36**, 1105-1121.

52. **Filiatrault, M.J., Stodghill, P.V., Bronstein, P.A., Moll, S., Lindeberg, M., Grills, G., Schweitzer, P., Wang, W., Schroth, G.P., Luo, S., Khrebtukova, I., Yang, Y., Thannhauser, T., Butcher, B.G., Cartinhour, S. and Schneider, D.J.** (2010) Transcriptome analysis of *Pseudomonas syringae* identifies new genes, noncoding RNAs, and antisense activity. *J. Bacteriol.*, **192**, 2359-2372.

53. **Filiatrault, M.J., Stodghill, P.V., Myers, C.R., Bronstein, P.A., Butcher, B.G., Lam, H., Grills, G., Schweitzer, P., Wang, W., Schneider, D.J. and Cartinhour, S.W.** (2011) Genome-wide identification of transcriptional start sites in the plant pathogen *Pseudomonas syringae* pv. *tomato* str. DC3000. *PLoS ONE*, **6**, e29335.

54. **Findeiß, S., Schmidtke, C., Stadler, P.F. and Bonas, U.** (2010) A novel family of plasmid-transferred anti-sense ncRNAs. *RNA Biol*, **7**, 120-124.

55. **Fozo, E.M., Hemm, M.R. and Storz, G.** (2008) Small toxic proteins and the antisense RNAs that repress them. *Microbiol. Mol. Biol. Rev.*, **72**, 579-589.

56. **Fozo, E.M., Kawano, M., Fontaine, F., Kaya, Y., Mendieta, K.S., Jones, K.L., Ocampo, A., Rudd, K.E. and Storz, G.** (2008) Repression of small toxic protein synthesis by the Sib and OhsC small RNAs. *Mol. Microbiol.*, **70**, 1076-1093.

57. **Fozo, E.M., Makarova, K.S., Shabalina, S.A., Yutin, N., Koonin, E.V. and Storz, G.** (2010) Abundance of type I toxin-antitoxin systems in bacteria: searches for new candidates and discovery of novel families. *Nucleic Acids Res.*, **38**, 3743-3759.

58. **Fransson, A., Ruusala, A. and Aspenstrom, P.** (2003) Atypical Rho GTPases have roles in mitochondrial homeostasis and apoptosis. *J. Biol. Chem.*, **278**, 6495-6502.

59. **Franze de Fernandez, M.T., Eoyang, L. and August, J.T.** (1968) Factor fraction required for the synthesis of bacteriophage Qbeta-RNA. *Nature*, **219**, 588-590.

60. **Fröhlich, K.S. and Vogel, J.** (2009) Activation of gene expression by small RNA. *Curr. Opin. Microbiol.*, **12**, 674-682.

61. **Geisinger, E., Adhikari, R.P., Jin, R., Ross, H.F. and Novick, R.P.** (2006) Inhibition of *rot* translation by RNAIII, a key feature of *agr* function. *Mol. Microbiol.*, **61**, 1038-1048.

62. **Geissmann, T., Chevalier, C., Cros, M.J., Boisset, S., Fechter, P., Noirot, C., Schrenzel, J., Francois, P., Vandenesch, F., Gaspin, C. and Romby, P.** (2009) A search for small noncoding RNAs in *Staphylococcus aureus* reveals a conserved sequence motif for regulation. *Nucleic Acids Res.*, **37**, 7239-7257.

63. **Genin, S.** (2010) Molecular traits controlling host range and adaptation to plants in *Ralstonia solanacearum*. *New Phytol.*, **187**, 920-928.

64. **Georg, J. and Hess, W.R.** (2011) *cis*-antisense RNA, another level of gene regulation in bacteria. *Microbiol. Mol. Biol. Rev.*, **75**, 286-300.

65. **Gerdes, K., Bech, F.W., Jorgensen, S.T., Lobner-Olesen, A., Rasmussen, P.B., Atlung, T., Boe, L., Karlstrom, O., Molin, S. and von Meyenburg, K.** (1986) Mechanism of postsegregational killing by the *hok* gene product of the *parB* system of plasmid R1 and its homology with the *relF* gene product of the *E. coli relB* operon. *EMBO J.*, **5**, 2023-2029.

66. **Gerdes, K. and Maisonneuve, E.** (2012) Bacterial persistence and toxin-antitoxin loci. *Annu. Rev. Microbiol.*, **66**, 103-123.

67. **Gottesman, S. and Storz, G.** (2011) Bacterial small RNA regulators: versatile roles and rapidly evolving variations. *Cold Spring Harb Perspect Biol*, **3**, doi: 10.1101/cshperspect.a003798.

68. **Greenberg, J.T. and Yao, N.** (2004) The role and regulation of programmed cell death in plant-pathogen interactions. *Cell. Microbiol.*, **6**, 201-211.

69. **Grundy, F.J., Moir, T.R., Haldeman, M.T. and Henkin, T.M.** (2002) Sequence requirements for terminators and antiterminators in the T box transcription antitermination system: disparity between conservation and functional requirements. *Nucleic Acids Res.*, **30**, 1646-1655.

70. **Gualerzi, C.O., Brandi, L., Caserta, E., Garofalo, C., Lammi, M., La Teana, A., Petrelli, D., Spurio, R., Tomsic, J. and Pon, C.L.** (2001) Initiation factors in the early events of mRNA translation in bacteria. *Cold Spring Harb. Symp. Quant. Biol.*, **66**, 363-376.

71. **Güell, M., van Noort, V., Yus, E., Chen, W.H., Leigh-Bell, J., Michalodimitrakis, K., Yamada, T., Arumugam, M., Doerks, T., Kuhner, S., Rode, M., Suyama, M., Schmidt, S., Gavin, A.C., Bork, P. and Serrano, L.** (2009) Transcriptome complexity in a genome-reduced bacterium. *Science*, **326**, 1268-1271.

72. **Guerrier-Takada, C., Gardiner, K., Marsh, T., Pace, N. and Altman, S.** (1983) The RNA moiety of ribonuclease P is the catalytic subunit of the enzyme. *Cell*, **35**, 849-857.

73. **Guo, Y., Figueiredo, F., Jones, J. and Wang, N.** (2011) HrpG and HrpX play global roles in coordinating different virulence traits of *Xanthomonas axonopodis* pv. *citri*. *Mol. Plant Microbe Interact.*, **24**, 649-661.

74. **Hajnsdorf, E. and Boni, I.V.** (2012) Multiple activities of RNA-binding proteins S1 and Hfq. *Biochimie*, **94**, 1544-1553.

75. **He, S.Y., Nomura, K. and Whittam, T.S.** (2004) Type III protein secretion mechanism in mammalian and plant pathogens. *Biochim. Biophys. Acta*, **1694**, 181-206.

76. **He, Y.W., Ng, A.Y., Xu, M., Lin, K., Wang, L.H., Dong, Y.H. and Zhang, L.H.** (2007) *Xanthomonas campestris* cell-cell communication involves a putative nucleotide receptor protein Clp and a hierarchical signalling network. *Mol. Microbiol.*, **64**, 281-292.

77. **He, Y.W., Xu, M., Lin, K., Ng, Y.J., Wen, C.M., Wang, L.H., Liu, Z.D., Zhang, H.B., Dong, Y.H., Dow, J.M. and Zhang, L.H.** (2006) Genome scale analysis of diffusible signal factor regulon in *Xanthomonas campestris* pv. *campestris*: identification of novel cell-cell communication-dependent genes and functions. *Mol. Microbiol.*, **59**, 610-622.

78. **Hengge, R.** (2009) Principles of c-di-GMP signalling in bacteria. *Nat. Rev. Microbiol.*, **7**, 263-273.
79. **Henkin, T.M.** (2008) Riboswitch RNAs: using RNA to sense cellular metabolism. *Genes Dev.*, **22**, 3383-3390.
80. **Herrick, J. and Sclavi, B.** (2007) Ribonucleotide reductase and the regulation of DNA replication: an old story and an ancient heritage. *Mol. Microbiol.*, **63**, 22-34.
81. **Herskovits, A.A., Bochkareva, E.S. and Bibi, E.** (2000) New prospects in studying the bacterial signal recognition particle pathway. *Mol. Microbiol.*, **38**, 927-939.
82. **Herzfeld, E.-M.** (2013), Halle (Saale), Universitäts-und Landesbibliothek Sachsen-Anhalt, Diss., 2013.
83. **Higgins, B.B.** (1922) The bacterial spot of pepper. *Phytopathology*, **12**, 501-517.
84. **Holmqvist, E., Reimegard, J., Sterk, M., Grantcharova, N., Römling, U. and Wagner, E.G.** (2010) Two antisense RNAs target the transcriptional regulator CsgD to inhibit curli synthesis. *EMBO J.*, **29**, 1840-1850.
85. **Hotson, A., Chosed, R., Shu, H., Orth, K. and Mudgett, M.B.** (2003) *Xanthomonas* type III effector XopD targets SUMO-conjugated proteins *in planta*. *Mol. Microbiol.*, **50**, 377-389.
86. **Hui, A. and de Boer, H.A.** (1987) Specialized ribosome system: preferential translation of a single mRNA species by a subpopulation of mutated ribosomes in *Escherichia coli. Proc. Natl. Acad. Sci. U.S.A.*, **84**, 4762-4766.
87. **Huntzinger, E., Boisset, S., Saveanu, C., Benito, Y., Geissmann, T., Namane, A., Lina, G., Etienne, J., Ehresmann, B., Ehresmann, C., Jacquier, A., Vandenesch, F. and Romby, P.** (2005) *Staphylococcus aureus* RNAIII and the endoribonuclease III coordinately regulate *spa* gene expression. *EMBO J.*, **24**, 824-835.
88. **Ikeda, Y., Yagi, M., Morita, T. and Aiba, H.** (2011) Hfq binding at RhlB-recognition region of RNase E is crucial for the rapid degradation of target mRNAs mediated by sRNAs in *Escherichia coli. Mol. Microbiol.*, **79**, 419-432.
89. **Jacob, F. and Monod, J.** (1961) Genetic regulatory mechanisms in the synthesis of proteins. *J. Mol. Biol.*, **3**, 318-356.
90. **Jacob, W.F., Santer, M. and Dahlberg, A.E.** (1987) A single base change in the Shine-Dalgarno region of 16S rRNA of *Escherichia coli* affects translation of many proteins. *Proc. Natl. Acad. Sci. U.S.A.*, **84**, 4757-4761.
91. **Jarrell, K.F. and McBride, M.J.** (2008) The surprisingly diverse ways that prokaryotes move. *Nat. Rev. Microbiol.*, **6**, 466-476.
92. **Jiang, R.P., Tang, D.J., Chen, X.L., He, Y.Q., Feng, J.X., Jiang, B.L., Lu, G.T., Lin, M. and Tang, J.L.** (2010) Identification of four novel small non-coding RNAs from *Xanthomonas campestris* pathovar *campestris*. *BMC Genomics*, **11**, 1-9.
93. **Johansson, J., Mandin, P., Renzoni, A., Chiaruttini, C., Springer, M. and Cossart, P.** (2002) An RNA thermosensor controls expression of virulence genes in *Listeria monocytogenes*. *Cell*, **110**, 551-561.
94. **Jones, J.B., Lacy, G.H., Bouzar, H., Stall, R.E. and Schaad, N.W.** (2004) Reclassification of the xanthomonads associated with bacterial spot disease of tomato and pepper. *Syst. Appl. Microbiol.*, **27**, 755-762.
95. **Jones, J.B., Stall, R.E. and Bouzar, H.** (1998) Diversity among xanthomonads pathogenic on pepper and tomato. *Annu. Rev. Phytopathol.*, **36**, 41-58.
96. **Jones, J.D. and Dangl, J.L.** (2006) The plant immune system. *Nature*, **444**, 323-329.
97. **Joyce, C.M. and Grindley, N.D.** (1982) Identification of two genes immediately downstream from the *polA* gene of *Escherichia coli. J. Bacteriol.*, **152**, 1211-1219.
98. **Kalamorz, F., Reichenbach, B., Marz, W., Rak, B. and Görke, B.** (2007) Feedback control of glucosamine-6-phosphate synthase GlmS expression depends on the small RNA GlmZ and involves the novel protein YhbJ in *Escherichia coli. Mol. Microbiol.*, **65**, 1518-1533.
99. **Kawano, M., Aravind, L. and Storz, G.** (2007) An antisense RNA controls synthesis of an SOS-induced toxin evolved from an antitoxin. *Mol. Microbiol.*, **64**, 738-754.
100. **Kawano, M., Reynolds, A.A., Miranda-Rios, J. and Storz, G.** (2005) Detection of 5'- and 3'-UTR-derived small RNAs and *cis*-encoded antisense RNAs in *Escherichia coli. Nucleic Acids Res.*, **33**, 1040-1050.
101. **Kim, J., Lee, S., Shin, H., Kim, S.C. and Cho, B.-K.** (2012) Elucidation of bacterial genome complexity using next-generation sequencing. *Biotechnol. Bioprocess Eng.*, **17**, 887-899.
102. **Kim, J.G., Stork, W. and Mudgett, M.B.** (2013) *Xanthomonas* type III effector XopD desumoylates tomato transcription factor SlERF4 to suppress ethylene responses and promote pathogen growth. *Cell Host Microbe*, **13**, 143-154.
103. **Kim, J.G., Taylor, K.W., Hotson, A., Keegan, M., Schmelz, E.A. and Mudgett, M.B.** (2008) XopD SUMO protease affects host transcription, promotes pathogen growth, and delays symptom development in *Xanthomonas*-infected tomato leaves. *Plant Cell*, **20**, 1915-1929.

104.   **Kirchberg, J., Büttner, D., Thiemer, B. and Sawers, R.G.** (2012) Aconitase B is required for optimal growth of *Xanthomonas campestris* pv. *vesicatoria* in pepper plants. *PLoS ONE*, **7**, e34941.

105.   **Koebnik, R., Krüger, A., Thieme, F., Urban, A. and Bonas, U.** (2006) Specific binding of the *Xanthomonas campestris* pv. vesicatoria AraC-type transcriptional activator HrpX to plant-inducible promoter boxes. *J. Bacteriol.*, **188**, 7652-7660.

106.   **Komarova, A.V., Tchufistova, L.S., Dreyfus, M. and Boni, I.V.** (2005) AU-rich sequences within 5' untranslated leaders enhance translation and stabilize mRNA in *Escherichia coli. J. Bacteriol.*, **187**, 1344-1349.

107.   **Kortmann, J. and Narberhaus, F.** (2012) Bacterial RNA thermometers: molecular zippers and switches. *Nat. Rev. Microbiol.*, **10**, 255-265.

108.   **Kröger, C., Dillon, S.C., Cameron, A.D., Papenfort, K., Sivasankaran, S.K., Hokamp, K., Chao, Y., Sittka, A., Hebrard, M., Handler, K., Colgan, A., Leekitcharoenphon, P., Langridge, G.C., Lohan, A.J., Loftus, B., Lucchini, S., Ussery, D.W., Dorman, C.J., Thomson, N.R., Vogel, J. and Hinton, J.C.** (2012) The transcriptional landscape and small RNAs of *Salmonella enterica* serovar Typhimurium. *Proc. Natl. Acad. Sci. U.S.A.*, **109**, E1277-1286.

109.   **Kunze, G., Zipfel, C., Robatzek, S., Niehaus, K., Boller, T. and Felix, G.** (2004) The N terminus of bacterial elongation factor Tu elicits innate immunity in *Arabidopsis* plants. *Plant Cell*, **16**, 3496-3507.

110.   **Lapouge, K., Schubert, M., Allain, F.H. and Haas, D.** (2008) Gac/Rsm signal transduction pathway of gamma-proteobacteria: from RNA recognition to regulation of social behaviour. *Mol. Microbiol.*, **67**, 241-253.

111.   **Laursen, B.S., Sørensen, H.P., Mortensen, K.K. and Sperling-Petersen, H.U.** (2005) Initiation of protein synthesis in bacteria. *Microbiol. Mol. Biol. Rev.*, **69**, 101-123.

112.   **Lee, B.M., Park, Y.J., Park, D.S., Kang, H.W., Kim, J.G., Song, E.S., Park, I.C., Yoon, U.H., Hahn, J.H., Koo, B.S., Lee, G.B., Kim, H., Park, H.S., Yoon, K.O., Kim, J.H., Jung, C.H., Koh, N.H., Seo, J.S. and Go, S.J.** (2005) The genome sequence of *Xanthomonas oryzae* pathovar *oryzae* KACC10331, the bacterial blight pathogen of rice. *Nucleic Acids Res.*, **33**, 577-586.

113.   **Lee, K., Huang, X., Yang, C., Lee, D., Ho, V., Nobuta, K., Fan, J.B. and Wang, K.** (2013) A genome-wide survey of highly expressed non-coding RNAs and biological validation of selected candidates in *Agrobacterium tumefaciens*. *PLoS ONE*, **8**, e70720.

114.   **Lenz, D.H., Mok, K.C., Lilley, B.N., Kulkarni, R.V., Wingreen, N.S. and Bassler, B.L.** (2004) The small RNA chaperone Hfq and multiple small RNAs control quorum sensing in *Vibrio harveyi* and *Vibrio cholerae. Cell*, **118**, 69-82.

115.   **Leyns, F., De Cleene, M., Swings, J.G. and De Ley, J.** (1984) The host range of the genus *Xanthomonas*. *The Botanical Review*, **50**, 308-356.

116.   **Li, R.F., Lu, G.T., Li, L., Su, H.Z., Feng, G.F., Chen, Y., He, Y.Q., Jiang, B.L., Tang, D.J. and Tang, J.L.** (2013) Identification of a putative cognate sensor kinase for the two-component response regulator HrpG, a key regulator controlling the expression of the *hrp* genes in *Xanthomonas campestris* pv. *campestris. Environ. Microbiol.*, doi: 10.1111/1462-2920.12207.

117.   **Li, W., Ying, X., Lu, Q. and Chen, L.** (2012) Predicting sRNAs and their targets in bacteria. *Genomics Proteomics Bioinformatics*, **10**, 276-284.

118.   **Liang, H., Zhao, Y.T., Zhang, J.Q., Wang, X.J., Fang, R.X. and Jia, Y.T.** (2011) Identification and functional characterization of small non-coding RNAs in *Xanthomonas oryzae* pathovar *oryzae. BMC Genomics*, **12**, 1-14.

119.   **Lioy, V.S., Martin, M.T., Camacho, A.G., Lurz, R., Antelmann, H., Hecker, M., Hitchin, E., Ridge, Y., Wells, J.M. and Alonso, J.C.** (2006) pSM19035-encoded zeta toxin induces stasis followed by death in a subpopulation of cells. *Microbiology*, **152**, 2365-2379.

120.   **Liu, M.Y., Gui, G., Wei, B., Preston, J.F., 3rd, Oakford, L., Yuksel, U., Giedroc, D.P. and Romeo, T.** (1997) The RNA molecule CsrB binds to the global regulatory protein CsrA and antagonizes its activity in *Escherichia coli. J. Biol. Chem.*, **272**, 17502-17510.

121.   **Liu, Y., Cui, Y., Mukherjee, A. and Chatterjee, A.K.** (1998) Characterization of a novel RNA regulator of *Erwinia carotovora* ssp. *carotovora* that controls production of extracellular enzymes and secondary metabolites. *Mol. Microbiol.*, **29**, 219-234.

122.   **Loh, E., Dussurget, O., Gripenland, J., Vaitkevicius, K., Tiensuu, T., Mandin, P., Repoila, F., Buchrieser, C., Cossart, P. and Johansson, J.r.** (2009) A *trans*-acting riboswitch controls expression of the virulence regulator PrfA in *Listeria monocytogenes. Cell*, **139**, 770-779.

123.   **Lu, X.H., An, S.Q., Tang, D.J., McCarthy, Y., Tang, J.L., Dow, J.M. and Ryan, R.P.** (2012) RsmA regulates biofilm formation in *Xanthomonas campestris* through a regulatory network involving cyclic di-GMP and the Clp transcription factor. *PLoS ONE*, **7**, e52646.

124.   **Ma, C. and Simons, R.W.** (1990) The IS10 antisense RNA blocks ribosome binding at the transposase translation initiation site. *EMBO J.*, **9**, 1267-1274.

125. **Majdalani, N., Cunning, C., Sledjeski, D., Elliott, T. and Gottesman, S.** (1998) DsrA RNA regulates translation of RpoS message by an anti-antisense mechanism, independent of its action as an antisilencer of transcription. *Proc. Natl. Acad. Sci. U.S.A.*, **95**, 12462-12467.

126. **Majdalani, N., Hernandez, D. and Gottesman, S.** (2002) Regulation and mode of action of the second small RNA activator of RpoS translation, RprA. *Mol. Microbiol.*, **46**, 813-826.

127. **Makarova, K.S., Wolf, Y.I. and Koonin, E.V.** (2009) Comprehensive comparative-genomic analysis of type 2 toxin-antitoxin systems and related mobile stress response systems in prokaryotes. *Biol. Direct*, **4**, 1-38.

128. **Malamud, F., Torres, P.S., Roeschlin, R., Rigano, L.A., Enrique, R., Bonomi, H.R., Castagnaro, A.P., Marano, M.R. and Vojnov, A.A.** (2011) The *Xanthomonas axonopodis* pv. *citri* flagellum is required for mature biofilm and canker development. *Microbiology*, **157**, 819-829.

129. **Mandal, M., Lee, M., Barrick, J.E., Weinberg, Z., Emilsson, G.M., Ruzzo, W.L. and Breaker, R.R.** (2004) A glycine-dependent riboswitch that uses cooperative binding to control gene expression. *Science*, **306**, 275-279.

130. **Mansfield, J., Genin, S., Magori, S., Citovsky, V., Sriariyanum, M., Ronald, P., Dow, M., Verdier, V., Beer, S.V., Machado, M.A., Toth, I., Salmond, G. and Foster, G.D.** (2012) Top 10 plant pathogenic bacteria in molecular plant pathology. *Mol. Plant Pathol.*, **13**, 614-629.

131. **Margulies, M., Egholm, M., Altman, W.E., Attiya, S., Bader, J.S., Bemben, L.A., Berka, J., Braverman, M.S., Chen, Y.J., Chen, Z., Dewell, S.B., Du, L., Fierro, J.M., Gomes, X.V., Godwin, B.C., He, W., Helgesen, S., Ho, C.H., Irzyk, G.P., Jando, S.C., Alenquer, M.L., Jarvie, T.P., Jirage, K.B., Kim, J.B., Knight, J.R., Lanza, J.R., Leamon, J.H., Lefkowitz, S.M., Lei, M., Li, J., Lohman, K.L., Lu, H., Makhijani, V.B., McDade, K.E., McKenna, M.P., Myers, E.W., Nickerson, E., Nobile, J.R., Plant, R., Puc, B.P., Ronan, M.T., Roth, G.T., Sarkis, G.J., Simons, J.F., Simpson, J.W., Srinivasan, M., Tartaro, K.R., Tomasz, A., Vogt, K.A., Volkmer, G.A., Wang, S.H., Wang, Y., Weiner, M.P., Yu, P., Begley, R.F. and Rothberg, J.M.** (2005) Genome sequencing in microfabricated high-density picolitre reactors. *Nature*, **437**, 376-380.

132. **Martin-Farmer, J. and Janssen, G.R.** (1999) A downstream CA repeat sequence increases translation from leadered and unleadered mRNA in *Escherichia coli*. *Mol. Microbiol.*, **31**, 1025-1038.

133. **McNealy, T.L., Forsbach-Birk, V., Shi, C. and Marre, R.** (2005) The Hfq homolog in *Legionella pneumophila* demonstrates regulation by LetA and RpoS and interacts with the global regulator CsrA. *J. Bacteriol.*, **187**, 1527-1532.

134. **Meyer, M.M., Hammond, M.C., Salinas, Y., Roth, A., Sudarsan, N. and Breaker, R.R.** (2011) Challenges of ligand identification for riboswitch candidates. *RNA Biol*, **8**, 5-10.

135. **Minsavage, G.V., Dahlbeck, D., Whalen, M.C., Kearny, B., Bonas, U., Staskawicz, B.J. and Stall, R.E.** (1990) Gene-for-gene relationships specifying disease resistance in *Xanthomonas campestris* pv. *vesicatoria* - pepper interactions. *Mol. Plant Microbe Interact.*, **3**, 41-47.

136. **Miranda-Rios, J., Navarro, M. and Soberon, M.** (2001) A conserved RNA structure (*thi* box) is involved in regulation of thiamin biosynthetic gene expression in bacteria. *Proc. Natl. Acad. Sci. U.S.A.*, **98**, 9736-9741.

137. **Mironov, A.S., Gusarov, I., Rafikov, R., Lopez, L.E., Shatalin, K., Kreneva, R.A., Perumov, D.A. and Nudler, E.** (2002) Sensing small molecules by nascent RNA: a mechanism to control transcription in bacteria. *Cell*, **111**, 747-756.

138. **Mitschke, J., Georg, J., Scholz, I., Sharma, C.M., Dienst, D., Bantscheff, J., Voss, B., Steglich, C., Wilde, A., Vogel, J. and Hess, W.R.** (2011) An experimentally anchored map of transcriptional start sites in the model cyanobacterium *Synechocystis* sp. PCC6803. *Proc. Natl. Acad. Sci. U.S.A.*, **108**, 2124-2129.

139. **Mizuno, T., Chou, M.Y. and Inouye, M.** (1984) A unique mechanism regulating gene expression: translational inhibition by a complementary RNA transcript (micRNA). *Proc. Natl. Acad. Sci. U.S.A.*, **81**, 1966-1970.

140. **Moll, I., Grill, S., Gualerzi, C.O. and Bläsi, U.** (2002) Leaderless mRNAs in bacteria: surprises in ribosomal recruitment and translational control. *Mol. Microbiol.*, **43**, 239-246.

141. **Møller, T., Franch, T., Udesen, C., Gerdes, K. and Valentin-Hansen, P.** (2002) Spot 42 RNA mediates discoordinate expression of the *E. coli* galactose operon. *Genes Dev.*, **16**, 1696-1706.

142. **Moore, S.D. and Sauer, R.T.** (2007) The tmRNA system for translational surveillance and ribosome rescue. *Annu. Rev. Biochem.*, **76**, 101-124.

143. **Morales, C.Q., Posada, J., Macneale, E., Franklin, D., Rivas, I., Bravo, M., Minsavage, J., Stall, R.E. and Whalen, M.C.** (2005) Functional analysis of the early chlorosis factor gene. *Mol. Plant Microbe Interact.*, **18**, 477-486.

144. **Morita, T., Maki, K. and Aiba, H.** (2005) RNase E-based ribonucleoprotein complexes: mechanical basis of mRNA destabilization mediated by bacterial noncoding RNAs. *Genes Dev.*, **19**, 2176-2186.

145.   **Morita, T., Mochizuki, Y. and Aiba, H.** (2006) Translational repression is sufficient for gene silencing by bacterial small noncoding RNAs in the absence of mRNA destruction. *Proc. Natl. Acad. Sci. U.S.A.*, **103**, 4858-4863.

146.   **Mraheil, M.A., Billion, A., Mohamed, W., Mukherjee, K., Kuenne, C., Pischimarov, J., Krawitz, C., Retey, J., Hartsch, T., Chakraborty, T. and Hain, T.** (2011) The intracellular sRNA transcriptome of *Listeria monocytogenes* during growth in macrophages. *Nucleic Acids Res.*, **39**, 4235-4248.

147.   **Nahvi, A., Sudarsan, N., Ebert, M.S., Zou, X., Brown, K.L. and Breaker, R.R.** (2002) Genetic control by a metabolite binding mRNA. *Chem. Biol.*, **9**, 1043-1049.

148.   **Nakagawa, S., Niimura, Y., Miura, K. and Gojobori, T.** (2010) Dynamic evolution of translation initiation mechanisms in prokaryotes. *Proc. Natl. Acad. Sci. U.S.A.*, **107**, 6382-6387.

149.   **Nocker, A., Hausherr, T., Balsiger, S., Krstulovic, N.P., Hennecke, H. and Narberhaus, F.** (2001) A mRNA-based thermosensor controls expression of rhizobial heat shock genes. *Nucleic Acids Res.*, **29**, 4800-4807.

150.   **Noël, L., Thieme, F., Nennstiel, D. and Bonas, U.** (2001) cDNA-AFLP analysis unravels a genome-wide *hrpG*-regulon in the plant pathogen *Xanthomonas campestris* pv. *vesicatoria*. *Mol. Microbiol.*, **41**, 1271-1281.

151.   **Novick, R.P., Ross, H.F., Projan, S.J., Kornblum, J., Kreiswirth, B. and Moghazeh, S.** (1993) Synthesis of staphylococcal virulence factors is controlled by a regulatory RNA molecule. *EMBO J.*, **12**, 3967-3975.

152.   **Nürnberger, T., Brunner, F., Kemmerling, B. and Piater, L.** (2004) Innate immunity in plants and animals: striking similarities and obvious differences. *Immunol. Rev.*, **198**, 249-266.

153.   **O'Donnell, S.M. and Janssen, G.R.** (2002) Leaderless mRNAs bind 70S ribosomes more strongly than 30S ribosomal subunits in *Escherichia coli*. *J. Bacteriol.*, **184**, 6730-6733.

154.   **Opdyke, J.A., Kang, J.G. and Storz, G.** (2004) GadY, a small-RNA regulator of acid response genes in *Escherichia coli*. *J. Bacteriol.*, **186**, 6698-6705.

155.   **Padalon-Brauch, G., Hershberg, R., Elgrably-Weiss, M., Baruch, K., Rosenshine, I., Margalit, H. and Altuvia, S.** (2008) Small RNAs encoded within genetic islands of *Salmonella typhimurium* show host-induced expression and role in virulence. *Nucleic Acids Res.*, **36**, 1913-1927.

156.   **Palmer, A.C., Ahlgren-Berg, A., Egan, J.B., Dodd, I.B. and Shearwin, K.E.** (2009) Potent transcriptional interference by pausing of RNA polymerases over a downstream promoter. *Mol. Cell*, **34**, 545-555.

157.   **Pandey, S.P., Minesinger, B.K., Kumar, J. and Walker, G.C.** (2011) A highly conserved protein of unknown function in *Sinorhizobium meliloti* affects sRNA regulation similar to Hfq. *Nucleic Acids Res.*, **39**, 4691-4708.

158.   **Papenfort, K., Pfeiffer, V., Lucchini, S., Sonawane, A., Hinton, J.C. and Vogel, J.** (2008) Systematic deletion of *Salmonella* small RNA genes identifies CyaR, a conserved CRP-dependent riboregulator of OmpX synthesis. *Mol. Microbiol.*, **68**, 890-906.

159.   **Papenfort, K., Pfeiffer, V., Mika, F., Lucchini, S., Hinton, J.C. and Vogel, J.** (2006) SigmaE-dependent small RNAs of *Salmonella* respond to membrane stress by accelerating global *omp* mRNA decay. *Mol. Microbiol.*, **62**, 1674-1688.

160.   **Papenfort, K. and Vogel, J.** (2010) Regulatory RNA in bacterial pathogens. *Cell Host Microbe*, **8**, 116-127.

161.   **Park, S.H., Butcher, B.G., Anderson, Z., Pellegrini, N., Bao, Z., D'Amico, K. and Filiatrault, M.J.** (2013) Analysis of the small RNA P16/RgsA in the plant pathogen *Pseudomonas syringae* pv. *tomato* strain DC3000. *Microbiology*, **159**, 296-306.

162.   **Perkins, T.T., Kingsley, R.A., Fookes, M.C., Gardner, P.P., James, K.D., Yu, L., Assefa, S.A., He, M., Croucher, N.J., Pickard, D.J., Maskell, D.J., Parkhill, J., Choudhary, J., Thomson, N.R. and Dougan, G.** (2009) A strand-specific RNA-Seq analysis of the transcriptome of the typhoid bacillus *Salmonella* Typhi. *PLoS Genet.*, **5**, e1000569.

163.   **Pfeiffer, V., Papenfort, K., Lucchini, S., Hinton, J.C. and Vogel, J.** (2009) Coding sequence targeting by MicC RNA reveals bacterial mRNA silencing downstream of translational initiation. *Nat. Struct. Mol. Biol.*, **16**, 840-846.

164.   **Pichon, C. and Felden, B.** (2008) Small RNA gene identification and mRNA target predictions in bacteria. *Bioinformatics*, **24**, 2807-2813.

165.   **Pieretti, I., Royer, M., Barbe, V., Carrere, S., Koebnik, R., Cociancich, S., Couloux, A., Darrasse, A., Gouzy, J., Jacques, M.A., Lauber, E., Manceau, C., Mangenot, S., Poussier, S., Segurens, B., Szurek, B., Verdier, V., Arlat, M. and Rott, P.** (2009) The complete genome sequence of *Xanthomonas albilineans* provides new insights into the reductive genome evolution of the xylem-limited *Xanthomonadaceae*. *BMC Genomics*, **10**, 1-15.

166.   **Pinto, A.C., Melo-Barbosa, H.P., Miyoshi, A., Silva, A. and Azevedo, V.** (2011) Application of RNA-seq to reveal the transcript profile in bacteria. *Genet. Mol. Res.*, **10**, 1707-1718.

167. **Pitzschke, A. and Hirt, H.** (2010) New insights into an old story: *Agrobacterium*-induced tumour formation in plants by plant transformation. *EMBO J.*, **29**, 1021-1032.

168. **Pohronezny, K., Moss, M.A., Dankers, W. and Schenk, J.** (1990) Dispersal and management of *Xanthomonas campestris* pv. *vesicatoria* during thinning of direct-seeded tomato. *Plant Dis.*, **74**, 800-805.

169. **Prévost, K., Desnoyers, G., Jacques, J.F., Lavoie, F. and Massé, E.** (2011) Small RNA-induced mRNA degradation achieved through both translation block and activated cleavage. *Genes Dev.*, **25**, 385-396.

170. **Qian, W., Jia, Y., Ren, S.X., He, Y.Q., Feng, J.X., Lu, L.F., Sun, Q., Ying, G., Tang, D.J., Tang, H., Wu, W., Hao, P., Wang, L., Jiang, B.L., Zeng, S., Gu, W.Y., Lu, G., Rong, L., Tian, Y., Yao, Z., Fu, G., Chen, B., Fang, R., Qiang, B., Chen, Z., Zhao, G.P., Tang, J.L. and He, C.** (2005) Comparative and functional genomic analyses of the pathogenicity of phytopathogen *Xanthomonas campestris* pv. *campestris*. *Genome Res.*, **15**, 757-767.

171. **Rajagopal, L., Sundari, C.S., Balasubramanian, D. and Sonti, R.V.** (1997) The bacterial pigment xanthomonadin offers protection against photodamage. *FEBS Lett.*, **415**, 125-128.

172. **Ray, S.K., Rajeshwari, R., Sharma, Y. and Sonti, R.V.** (2002) A high-molecular-weight outer membrane protein of *Xanthomonas oryzae* pv. *oryzae* exhibits similarity to non-fimbrial adhesins of animal pathogenic bacteria and is required for optimum virulence. *Mol. Microbiol.*, **46**, 637-647.

173. **Rice, P.W. and Dahlberg, J.E.** (1982) A gene between *polA* and *glnA* retards growth of *Escherichia coli* when present in multiple copies: physiological effects of the gene for spot 42 RNA. *J. Bacteriol.*, **152**, 1196-1210.

174. **Richards, G.R. and Vanderpool, C.K.** (2011) Molecular call and response: the physiology of bacterial small RNAs. *Biochim. Biophys. Acta*, **1809**, 525-531.

175. **Ringquist, S., MacDonald, M., Gibson, T. and Gold, L.** (1993) Nature of the ribosomal mRNA track: analysis of ribosome-binding sites containing different sequences and secondary structures. *Biochemistry*, **32**, 10254-10262.

176. **Roberts, M.W. and Rabinowitz, J.C.** (1989) The effect of *Escherichia coli* ribosomal protein S1 on the translational specificity of bacterial ribosomes. *J. Biol. Chem.*, **264**, 2228-2235.

177. **Romeo, T., Gong, M., Liu, M.Y. and Brun-Zinkernagel, A.M.** (1993) Identification and molecular characterization of *csrA*, a pleiotropic gene from *Escherichia coli* that affects glycogen biosynthesis, gluconeogenesis, cell size, and surface properties. *J. Bacteriol.*, **175**, 4744-4755.

178. **Ronald, P.C. and Staskawicz, B.J.** (1988) The avirulence gene *avrBs1* from *Xanthomonas campestris* pv. *vesicatoria* encodes a 50-kD protein. *Mol. Plant Microbe Interact.*, **1**, 191-198.

179. **Roscetto, E., Angrisano, T., Costa, V., Casalino, M., Forstner, K.U., Sharma, C.M., Di Nocera, P.P. and De Gregorio, E.** (2012) Functional characterization of the RNA chaperone Hfq in the opportunistic human pathogen *Stenotrophomonas maltophilia*. *J. Bacteriol.*, **194**, 5864-5874.

180. **Rossier, O., Wengelnik, K., Hahn, K. and Bonas, U.** (1999) The *Xanthomonas* Hrp type III system secretes proteins from plant and mammalian bacterial pathogens. *Proc. Natl. Acad. Sci. U.S.A.*, **96**, 9368-9373.

181. **Rothberg, J.M., Hinz, W., Rearick, T.M., Schultz, J., Mileski, W., Davey, M., Leamon, J.H., Johnson, K., Milgrew, M.J., Edwards, M., Hoon, J., Simons, J.F., Marran, D., Myers, J.W., Davidson, J.F., Branting, A., Nobile, J.R., Puc, B.P., Light, D., Clark, T.A., Huber, M., Branciforte, J.T., Stoner, I.B., Cawley, S.E., Lyons, M., Fu, Y., Homer, N., Sedova, M., Miao, X., Reed, B., Sabina, J., Feierstein, E., Schorn, M., Alanjary, M., Dimalanta, E., Dressman, D., Kasinskas, R., Sokolsky, T., Fidanza, J.A., Namsaraev, E., McKernan, K.J., Williams, A., Roth, G.T. and Bustillo, J.** (2011) An integrated semiconductor device enabling non-optical genome sequencing. *Nature*, **475**, 348-352.

182. **Ryan, R.P., Fouhy, Y., Lucey, J.F., Jiang, B.L., He, Y.Q., Feng, J.X., Tang, J.L. and Dow, J.M.** (2007) Cyclic di-GMP signalling in the virulence and environmental adaptation of *Xanthomonas campestris*. *Mol. Microbiol.*, **63**, 429-442.

183. **Sahagan, B.G. and Dahlberg, J.E.** (1979) A small, unstable RNA molecule of *Escherichia coli*: spot 42 RNA. I. Nucleotide sequence analysis. *J. Mol. Biol.*, **131**, 573-592.

184. **Said, N., Rieder, R., Hurwitz, R., Deckert, J., Urlaub, H. and Vogel, J.** (2009) *In vivo* expression and purification of aptamer-tagged small RNA regulators. *Nucleic Acids Res.*, **37**, e133.

185. **Salah, P., Bisaglia, M., Aliprandi, P., Uzan, M., Sizun, C. and Bontems, F.** (2009) Probing the relationship between Gram-negative and Gram-positive S1 proteins by sequence analysis. *Nucleic Acids Res.*, **37**, 5578-5588.

186. **Schlüter, J.P., Reinkensmeier, J., Daschkey, S., Evguenieva-Hackenberg, E., Janssen, S., Jänicke, S., Becker, J.D., Giegerich, R. and Becker, A.** (2010) A genome-wide survey of sRNAs in the symbiotic nitrogen-fixing alpha-proteobacterium *Sinorhizobium meliloti*. *BMC Genomics*, **11**, 1-35.

187. **Schmidtke, C., Abendroth, U., Brock, J., Serrania, J., Becker, A. and Bonas, U.** (2013) Small RNA sX13: a multifaceted regulator of virulence in the plant pathogen *Xanthomonas*. *PLoS Pathog.*, **9**, e1003626.

188. **Schmidtke, C., Findeiß, S., Sharma, C.M., Kuhfuss, J., Hoffmann, S., Vogel, J., Stadler, P.F. and Bonas, U.** (2012) Genome-wide transcriptome analysis of the plant pathogen *Xanthomonas* identifies sRNAs with putative virulence functions. *Nucleic Acids Res.*, **40**, 2020-2031.

189. **Schulte, R. and Bonas, U.** (1992) Expression of the *Xanthomonas campestris* pv. vesicatoria *hrp* gene cluster, which determines pathogenicity and hypersensitivity on pepper and tomato, is plant inducible. *J. Bacteriol.*, **174**, 815-823.

190. **Schulze, S.** (2013), Halle (Saale), Universitäts-und Landesbibliothek Sachsen-Anhalt, Diss., 2013.

191. **Schulze, S., Kay, S., Büttner, D., Egler, M., Eschen-Lippold, L., Hause, G., Krüger, A., Lee, J., Müller, O., Scheel, D., Szczesny, R., Thieme, F. and Bonas, U.** (2012) Analysis of new type III effectors from *Xanthomonas* uncovers XopB and XopS as suppressors of plant immunity. *New Phytol.*, **195**, 894-911.

192. **Schuster, C.F. and Bertram, R.** (2013) Toxin-antitoxin systems are ubiquitous and versatile modulators of prokaryotic cell fate. *FEMS Microbiol. Lett.*, **340**, 73-85.

193. **Serganov, A. and Nudler, E.** (2013) A decade of riboswitches. *Cell*, **152**, 17-24.

194. **Serganov, A. and Patel, D.J.** (2009) Amino acid recognition and gene regulation by riboswitches. *Biochim. Biophys. Acta*, **1789**, 592-611.

195. **Sharma, C.M., Darfeuille, F., Plantinga, T.H. and Vogel, J.** (2007) A small RNA regulates multiple ABC transporter mRNAs by targeting C/A-rich elements inside and upstream of ribosome-binding sites. *Genes Dev.*, **21**, 2804-2817.

196. **Sharma, C.M., Hoffmann, S., Darfeuille, F., Reignier, J., Findeiss, S., Sittka, A., Chabas, S., Reiche, K., Hackermüller, J., Reinhardt, R., Stadler, P.F. and Vogel, J.** (2010) The primary transcriptome of the major human pathogen *Helicobacter pylori*. *Nature*, **464**, 250-255.

197. **Sharma, C.M., Papenfort, K., Pernitzsch, S.R., Mollenkopf, H.J., Hinton, J.C. and Vogel, J.** (2011) Pervasive post-transcriptional control of genes involved in amino acid metabolism by the Hfq-dependent GcvB small RNA. *Mol. Microbiol.*, **81**, 1144-1165.

198. **Sharma, C.M. and Vogel, J.** (2009) Experimental approaches for the discovery and characterization of regulatory small RNA. *Curr. Opin. Microbiol.*, **12**, 536-546.

199. **Shendure, J., Porreca, G.J., Reppas, N.B., Lin, X., McCutcheon, J.P., Rosenbaum, A.M., Wang, M.D., Zhang, K., Mitra, R.D. and Church, G.M.** (2005) Accurate multiplex polony sequencing of an evolved bacterial genome. *Science*, **309**, 1728-1732.

200. **Shine, J. and Dalgarno, L.** (1974) The 3'-terminal sequence of *Escherichia coli* 16S ribosomal RNA: complementarity to nonsense triplets and ribosome binding sites. *Proc. Natl. Acad. Sci. U.S.A.*, **71**, 1342-1346.

201. **Simons, R.W. and Kleckner, N.** (1983) Translational control of IS10 transposition. *Cell*, **34**, 683-691.

202. **Sittka, A., Lucchini, S., Papenfort, K., Sharma, C.M., Rolle, K., Binnewies, T.T., Hinton, J.C. and Vogel, J.** (2008) Deep sequencing analysis of small noncoding RNA and mRNA targets of the global post-transcriptional regulator, Hfq. *PLoS Genet.*, **4**, e1000163.

203. **Soper, T., Mandin, P., Majdalani, N., Gottesman, S. and Woodson, S.A.** (2010) Positive regulation by small RNAs and the role of Hfq. *Proc. Natl. Acad. Sci. U.S.A.*, **107**, 9602-9607.

204. **Sorek, R. and Cossart, P.** (2010) Prokaryotic transcriptomics: a new view on regulation, physiology and pathogenicity. *Nat. Rev. Genet.*, **11**, 9-16.

205. **Sorek, R., Lawrence, C.M. and Wiedenheft, B.** (2013) CRISPR-mediated adaptive immune systems in bacteria and archaea. *Annu. Rev. Biochem.*, **82**, 237-266.

206. **Stall, R.E., Jones, J.B. and Minsavage, G.V.** (2009) Durability of resistance in tomato and pepper to xanthomonads causing bacterial spot. *Annu. Rev. Phytopathol.*, **47**, 265-284.

207. **Stazic, D., Lindell, D. and Steglich, C.** (2011) Antisense RNA protects mRNA from RNase E degradation by RNA-RNA duplex formation during phage infection. *Nucleic Acids Res.*, **39**, 4890-4899.

208. **Steitz, J.A. and Jakes, K.** (1975) How ribosomes select initiator regions in mRNA: base pair formation between the 3' terminus of 16S rRNA and the mRNA during initiation of protein synthesis in *Escherichia coli*. *Proc. Natl. Acad. Sci. U.S.A.*, **72**, 4734-4738.

209. **Stork, M., Di Lorenzo, M., Welch, T.J. and Crosa, J.H.** (2007) Transcription termination within the iron transport-biosynthesis operon of *Vibrio anguillarum* requires an antisense RNA. *J. Bacteriol.*, **189**, 3479-3488.

210. **Storz, G., Vogel, J. and Wassarman, K.M.** (2011) Regulation by small RNAs in bacteria: expanding frontiers. *Mol. Cell*, **43**, 880-891.

211. **Stougaard, P., Molin, S. and Nordström, K.** (1981) RNAs involved in copy-number control and incompatibility of plasmid R1. *Proc. Natl. Acad. Sci. U.S.A.*, **78**, 6008-6012.

212. **Swings, J.-G. and Civerolo, E.L.** (1993) *Xanthomonas*. Chapman & Hall, London, UK.

213.   **Szczesny, R., Jordan, M., Schramm, C., Schulz, S., Cogez, V., Bonas, U. and Büttner, D.** (2010) Functional characterization of the Xcs and Xps type II secretion systems from the plant pathogenic bacterium *Xanthomonas campestris* pv. *vesicatoria*. *New Phytol.*, **187**, 983-1002.

214.   **Tang, J.L., Liu, Y.N., Barber, C.E., Dow, J.M., Wootton, J.C. and Daniels, M.J.** (1991) Genetic and molecular analysis of a cluster of *rpf* genes involved in positive regulation of synthesis of extracellular enzymes and polysaccharide in *Xanthomonas campestris* pathovar *campestris*. *Mol. Gen. Genet.*, **226**, 409-417.

215.   **Thieme, F., Koebnik, R., Bekel, T., Berger, C., Boch, J., Büttner, D., Caldana, C., Gaigalat, L., Goesmann, A., Kay, S., Kirchner, O., Lanz, C., Linke, B., McHardy, A.C., Meyer, F., Mittenhuber, G., Nies, D.H., Niesbach-Klösgen, U., Patschkowski, T., Rückert, C., Rupp, O., Schneiker, S., Schuster, S.C., Vorhölter, F.J., Weber, E., Pühler, A., Bonas, U., Bartels, D. and Kaiser, O.** (2005) Insights into genome plasticity and pathogenicity of the plant pathogenic bacterium *Xanthomonas campestris* pv. vesicatoria revealed by the complete genome sequence. *J. Bacteriol.*, **187**, 7254-7266.

216.   **Thieme, F., Szczesny, R., Urban, A., Kirchner, O., Hause, G. and Bonas, U.** (2007) New type III effectors from *Xanthomonas campestris* pv. *vesicatoria* trigger plant reactions dependent on a conserved N-myristoylation motif. *Mol. Plant Microbe Interact.*, **20**, 1250-1261.

217.   **Timmermans, J. and Van Melderen, L.** (2010) Post-transcriptional global regulation by CsrA in bacteria. *Cell. Mol. Life Sci.*, **67**, 2897-2908.

218.   **Toledo-Arana, A., Dussurget, O., Nikitas, G., Sesto, N., Guet-Revillet, H., Balestrino, D., Loh, E., Gripenland, J., Tiensuu, T., Vaitkevicius, K., Barthelemy, M., Vergassola, M., Nahori, M.A., Soubigou, G., Regnault, B., Coppee, J.Y., Lecuit, M., Johansson, J. and Cossart, P.** (2009) The *Listeria* transcriptional landscape from saprophytism to virulence. *Nature*, **459**, 950-956.

219.   **Tomizawa, J., Itoh, T., Selzer, G. and Som, T.** (1981) Inhibition of ColE1 RNA primer formation by a plasmid-specified small RNA. *Proc. Natl. Acad. Sci. U.S.A.*, **78**, 1421-1425.

220.   **Tramonti, A., De Canio, M. and De Biase, D.** (2008) GadX/GadW-dependent regulation of the *Escherichia coli* acid fitness island: transcriptional control at the *gadY-gadW* divergent promoters and identification of four novel 42 bp GadX/GadW-specific binding sites. *Mol. Microbiol.*, **70**, 965-982.

221.   **Trotochaud, A.E. and Wassarman, K.M.** (2004) 6S RNA function enhances long-term cell survival. *J. Bacteriol.*, **186**, 4978-4985.

222.   **Trotochaud, A.E. and Wassarman, K.M.** (2005) A highly conserved 6S RNA structure is required for regulation of transcription. *Nat. Struct. Mol. Biol.*, **12**, 313-319.

223.   **Udagawa, T., Shimizu, Y. and Ueda, T.** (2004) Evidence for the translation initiation of leaderless mRNAs by the intact 70 S ribosome without its dissociation into subunits in eubacteria. *J. Biol. Chem.*, **279**, 8539-8546.

224.   **Urban, J.H. and Vogel, J.** (2007) Translational control and target recognition by *Escherichia coli* small RNAs *in vivo*. *Nucleic Acids Res.*, **35**, 1018-1037.

225.   **van Doorn, J., Boonekamp, P.M. and Oudega, B.** (1994) Partial characterization of fimbriae of *Xanthomonas campestris* pv. *hyacinthi*. *Mol. Plant Microbe Interact.*, **7**, 334-344.

226.   **Vanderpool, C.K., Balasubramanian, D. and Lloyd, C.R.** (2011) Dual-function RNA regulators in bacteria. *Biochimie*, **93**, 1943-1949.

227.   **Vanderpool, C.K. and Gottesman, S.** (2004) Involvement of a novel transcriptional activator and small RNA in post-transcriptional regulation of the glucose phosphoenolpyruvate phosphotransferase system. *Mol. Microbiol.*, **54**, 1076-1089.

228.   **Vanneste, J.L.L.** (2000) Fire blight: the disease and its causative agent, *Erwinia amylovora*. CAB International, New York, USA.

229.   **Vauterin, L., Rademaker, J. and Swings, J.** (2000) Synopsis on the taxonomy of the genus *Xanthomonas*. *Phytopathology*, **90**, 677-682.

230.   **Viegas, S.C., Pfeiffer, V., Sittka, A., Silva, I.J., Vogel, J. and Arraiano, C.M.** (2007) Characterization of the role of ribonucleases in *Salmonella* small RNA decay. *Nucleic Acids Res.*, **35**, 7651-7664.

231.   **Vitreschak, A.G., Rodionov, D.A., Mironov, A.A. and Gelfand, M.S.** (2002) Regulation of riboflavin biosynthesis and transport genes in bacteria by transcriptional and translational attenuation. *Nucleic Acids Res.*, **30**, 3141-3151.

232.   **Vitreschak, A.G., Rodionov, D.A., Mironov, A.A. and Gelfand, M.S.** (2003) Regulation of the vitamin B_{12} metabolism and transport in bacteria by a conserved RNA structural element. *RNA*, **9**, 1084-1097.

233.   **Vogel, J., Argaman, L., Wagner, E.G. and Altuvia, S.** (2004) The small RNA IstR inhibits synthesis of an SOS-induced toxic peptide. *Curr. Biol.*, **14**, 2271-2276.

234.   **Vogel, J., Bartels, V., Tang, T.H., Churakov, G., Slagter-Jäger, J.G., Hüttenhofer, A. and Wagner, E.G.** (2003) RNomics in *Escherichia coli* detects new sRNA species and indicates parallel transcriptional output in bacteria. *Nucleic Acids Res.*, **31**, 6435-6443.

235.   **Vogel, J. and Luisi, B.F.** (2011) Hfq and its constellation of RNA. *Nat. Rev. Microbiol.*, **9**, 578-589.

236.   **Vogel, J. and Wagner, E.G.** (2007) Target identification of small noncoding RNAs in bacteria. *Curr. Opin. Microbiol.*, **10**, 262-270.

237.   **Wadler, C.S. and Vanderpool, C.K.** (2007) A dual function for a bacterial small RNA: SgrS performs base pairing-dependent regulation and encodes a functional polypeptide. *Proc. Natl. Acad. Sci. U.S.A.*, **104**, 20454-20459.

238.   **Waldminghaus, T., Heidrich, N., Brantl, S. and Narberhaus, F.** (2007) FourU: a novel type of RNA thermometer in *Salmonella*. *Mol. Microbiol.*, **65**, 413-424.

239.   **Wang, J.X., Lee, E.R., Morales, D.R., Lim, J. and Breaker, R.R.** (2008) Riboswitches that sense S-adenosylhomocysteine and activate genes involved in coenzyme recycling. *Mol. Cell*, **29**, 691-702.

240.   **Wang, L., Makino, S., Subedee, A. and Bogdanove, A.J.** (2007) Novel candidate virulence factors in rice pathogen *Xanthomonas oryzae* pv. *oryzicola* as revealed by mutational analysis. *Appl. Environ. Microbiol.*, **73**, 8023-8027.

241.   **Wang, L.H., He, Y., Gao, Y., Wu, J.E., Dong, Y.H., He, C., Wang, S.X., Weng, L.X., Xu, J.L., Tay, L., Fang, R.X. and Zhang, L.H.** (2004) A bacterial cell-cell communication signal with cross-kingdom structural analogues. *Mol. Microbiol.*, **51**, 903-912.

242.   **Wassarman, K.M.** (2007) 6S RNA: a small RNA regulator of transcription. *Curr. Opin. Microbiol.*, **10**, 164-168.

243.   **Wassarman, K.M., Repoila, F., Rosenow, C., Storz, G. and Gottesman, S.** (2001) Identification of novel small RNAs using comparative genomics and microarrays. *Genes Dev.*, **15**, 1637-1651.

244.   **Wassarman, K.M. and Storz, G.** (2000) 6S RNA regulates *E. coli* RNA polymerase activity. *Cell*, **101**, 613-623.

245.   **Waters, L.S. and Storz, G.** (2009) Regulatory RNAs in bacteria. *Cell*, **136**, 615-628.

246.   **Weber, E., Berger, C., Bonas, U. and Koebnik, R.** (2007) Refinement of the *Xanthomonas campestris* pv. *vesicatoria hrpD* and *hrpE* operon structure. *Mol. Plant Microbe Interact.*, **20**, 559-567.

247.   **Weber, E., Ojanen-Reuhs, T., Huguet, E., Hause, G., Romantschuk, M., Korhonen, T.K., Bonas, U. and Koebnik, R.** (2005) The type III-dependent Hrp pilus is required for productive interaction of *Xanthomonas campestris* pv. vesicatoria with pepper host plants. *J. Bacteriol.*, **187**, 2458-2468.

248.   **Wei, B.L., Brun-Zinkernagel, A.M., Simecka, J.W., Pruss, B.M., Babitzke, P. and Romeo, T.** (2001) Positive regulation of motility and *flhDC* expression by the RNA-binding protein CsrA of *Escherichia coli*. *Mol. Microbiol.*, **40**, 245-256.

249.   **Weiberg, A., Wang, M., Lin, F.M., Zhao, H., Zhang, Z., Kaloshian, I., Huang, H.D. and Jin, H.** (2013) Fungal small RNAs suppress plant immunity by hijacking host RNA interference pathways. *Science*, **342**, 118-123.

250.   **Weilbacher, T., Suzuki, K., Dubey, A.K., Wang, X., Gudapaty, S., Morozov, I., Baker, C.S., Georgellis, D., Babitzke, P. and Romeo, T.** (2003) A novel sRNA component of the carbon storage regulatory system of *Escherichia coli*. *Mol. Microbiol.*, **48**, 657-670.

251.   **Weissenmayer, B.A., Prendergast, J.G., Lohan, A.J. and Loftus, B.J.** (2011) Sequencing illustrates the transcriptional response of *Legionella pneumophila* during infection and identifies seventy novel small non-coding RNAs. *PLoS ONE*, **6**, e17570.

252.   **Wengelnik, K. and Bonas, U.** (1996) HrpXv, an AraC-type regulator, activates expression of five of the six loci in the *hrp* cluster of *Xanthomonas campestris* pv. vesicatoria. *J. Bacteriol.*, **178**, 3462-3469.

253.   **Wengelnik, K., Marie, C., Russel, M. and Bonas, U.** (1996) Expression and localization of HrpA1, a protein of *Xanthomonas campestris* pv. vesicatoria essential for pathogenicity and induction of the hypersensitive reaction. *J. Bacteriol.*, **178**, 1061-1069.

254.   **Wengelnik, K., Rossier, O. and Bonas, U.** (1999) Mutations in the regulatory gene *hrpG* of *Xanthomonas campestris* pv. vesicatoria result in constitutive expression of all *hrp* genes. *J. Bacteriol.*, **181**, 6828-6831.

255.   **Wengelnik, K., Van den Ackerveken, G. and Bonas, U.** (1996) HrpG, a key *hrp* regulatory protein of *Xanthomonas campestris* pv. *vesicatoria* is homologous to two-component response regulators. *Mol. Plant Microbe Interact.*, **9**, 704-712.

256.   **Westermann, A.J., Gorski, S.A. and Vogel, J.** (2012) Dual RNA-seq of pathogen and host. *Nat. Rev. Microbiol.*, **10**, 618-630.

257.   **Wilms, I., Möller, P., Stock, A.M., Gurski, R., Lai, E.M. and Narberhaus, F.** (2012) Hfq influences multiple transport systems and virulence in the plant pathogen *Agrobacterium tumefaciens*. *J. Bacteriol.*, **194**, 5209-5217.

258.   **Wilms, I., Overlöper, A., Nowrousian, M., Sharma, C.M. and Narberhaus, F.** (2012) Deep sequencing uncovers numerous small RNAs on all four replicons of the plant pathogen *Agrobacterium tumefaciens*. *RNA Biol*, **9**, 446-457.

259. **Wilms, I., Voss, B., Hess, W.R., Leichert, L.I. and Narberhaus, F.** (2011) Small RNA-mediated control of the *Agrobacterium tumefaciens* GABA binding protein. *Mol. Microbiol.*, **80**, 492-506.

260. **Winkler, W., Nahvi, A. and Breaker, R.R.** (2002) Thiamine derivatives bind messenger RNAs directly to regulate bacterial gene expression. *Nature*, **419**, 952-956.

261. **Yanofsky, C.** (1981) Attenuation in the control of expression of bacterial operons. *Nature*, **289**, 751-758.

262. **Zeng, Q., McNally, R.R. and Sundin, G.W.** (2013) Global small RNA chaperone Hfq and regulatory small RNAs are important virulence regulators in *Erwinia amylovora*. *J. Bacteriol.*, **195**, 1706-1717.

263. **Zhang, Y., Wei, C., Jiang, W., Wang, L., Li, C., Wang, Y., Dow, J.M. and Sun, W.** (2013) The HD-GYP domain protein RpfG of *Xanthomonas oryzae* pv. *oryzicola* regulates synthesis of extracellular polysaccharides that contribute to biofilm formation and virulence on rice. *PLoS ONE*, **8**, e59428.

264. **Zhu, P.L., Zhao, S., Tang, J.L. and Feng, J.X.** (2010) The *rsmA*-like gene $rsmA_{Xoo}$ of *Xanthomonas oryzae* pv. *oryzae* regulates bacterial virulence and production of diffusible signal factor. *Mol. Plant Pathol.*, **12**, 227-237.

265. **Zielenkiewicz, U. and Ceglowski, P.** (2005) The toxin-antitoxin system of the streptococcal plasmid pSM19035. *J. Bacteriol.*, **187**, 6094-6105.

Anhang zu Kapitel 2.1.1.: Tabellen S1 bis S9

Table S1. Bacterial strains, plasmids and oligonucleotides used in this study.

Strain or plasmid	Relevant characteristics[a]	Reference or source
Xanthomonas campestris* pv. *vesicatoria		
85-10	Pepper-race 2; wild type; Rif[R]	(1)
85-10Δ*hrpX*	85-10 derivative deleted in *hrpX*; Rif[R]	(2)
85-10Δ*sX12*	85-10 derivative deleted in *sX12*; Rif[R]	This study
85*	85-10 derivative containing the *hrpG* * mutation; Rif[R]	(3)
Escherichia coli		
DH5α λpir	F[-] *recA hsdR17(r$_k^-$,m$_k^+$)* φ80d*lacZ ΔM15 [λpir]*	(4)
TOP10	F[-] *mcrA* Δ(*mrr-hsd*RMS-*mcr*BC) φ80*lacZ*ΔM15 Δ*lac*X74 *rec*A1 *ara*D139 Δ(*ara-leu*) 7697 *gal*U *gal*K *rps*L (Str[R]) *end*A1 *nup*G	Invitrogen
TOP10F'	F'{*lacI*[q] Tn*10* (Tet[R]-)} *mcrA* Δ(*mrr-hsd*RMS-*mcr*BC) φ80*lacZ*ΔM15Δ*lac*X74 *rec*A1 *ara*D139 Δ(*ara-leu*) 7697 *gal*U *gal*K *rps*L *end*A1 *nup*G	Invitrogen
Plasmids		
pBRM-P	Derivative of pBBR1MCS-5 without promoter; Gm[R]	(5)
pBRM-sX6	pBRM-P derivative expressing c-Myc-epitope tagged sX6; Gm[R]	This study
pCR 2.1-TOPO	General-purpose cloning vector; Ap[R] Km[R]	Invitrogen
pFG72-1	Derivative of pUFR043 containing *hrpG**; Km[R]	(3)
pLAFR6	RK2 replicon, Mob[+] Tra[-]; without promoter; multicloning site flanked by transcription terminators; Tc[R]	(6)
p*sX12*	pLAFR6 derivative containing *sX12*; Tc[R]	This study
pOK1	Suicide vector; *sacB sacQ mobRK2* oriR6K; Sm[R]	(7)
pOKΔ*sX12*	pOK1 derivative containing flanking regions of *sX12*; Sm[R]	This study
pRK2013	ColE1 replicon, TraRK[+] Mob[+]; Km[R]	(8)
Oligonucleotide	**Sequence (Purpose)**	**Gene**
NB76	CTGCGTGGAGTTTCTAGGCT (Northern blot probe)	*sX1*
NB2	GCTGCCTAGATGCTCTAGGG (Northern blot probe)	*sX2*
NB12	GTGGCGAGTAAGGCAAAAAG (Northern blot probe)	*sX3*
NB13	CAAATTGTTCAGGAACTTACGC (Northern blot probe)	*sX4*
NB81	GGACTTCGTAACGCAGGACT (Northern blot probe)	*sX5*
NB16	TCAGGATGTGCTTGAACTTCAT (Northern blot probe)	*sX6*
NB17	CTCCTGGCCTTCGATAGATCTA (Northern blot probe)	*sX7*
NB23	ACTAGGCGCATAAGTCGTTGTT (Northern blot probe)	*sX8*
NB25	ACATCACACCGCGACCAG (Northern blot probe)	*sX9*
NB28	CCGAAGGAGTGTTGTATTAGC (Northern blot probe)	*sX10*
NB31	AGTAGTAGGCTTGCTCAGAGCC (Northern blot probe)	*sX11*
NB42	TACCTTTCGACGAGGATGTG (Northern blot probe)	*sX12*
NB100	GAGAGATCTACCGACAGGCGT (Northern blot probe)	*sX13*
NB101	CACAGCTTCCGACTGACATC (Northern blot probe)	*sX14*
NB67	TTACCGATCGTCGTGTAGCTG (Northern blot probe)	*sX15*
NB30	TGGTCGTAATTAAGGGACAAG (Northern blot probe)	*6S*
NB55	CTACCGACCCTTACGCTACC (Northern blot probe)	*asX1*
NB89	CAGACCACCCAACAACCTTC (Northern blot probe)	*asX2*
NB45	GCCGAACTTACGTCCTTGTC (Northern blot probe)	*asX3*
NB98	CCAAAAACACGACTCAGTCG (Northern blot probe)	*asX4*
NB37	GATCAAACGAGTGGCTACTGTG (Northern blot probe)	*asX5*
NB66	AAACAGACCATGTGGCACATC (Northern blot probe)	*asX6*
NB72	CAGCTACACGACGATCGGTA (Northern blot probe)	*asX7*
NB-5S	CCTGGCGATGACCTACTCTC (Northern blot probe)	*5S rRNA*
5'_RNA-adapter	AUAUGCGCGAAUUCCUGUAGAACGAACACUAGAAGAAA (5' RACE; RNA-adapter)	
3'_RNA-adapter*	P-UUCACUGUUCUUAGCGGCCGCAUGCUC-idT* (3' RACE; RNA-adapter)	
5'-RACE-adapter	GCGCGAATTCCTGTAGA (5' RACE; adapter-specific primer)	
3'-RACE-adapter	CGGCCGCTAAGAACAG (3' RACE; adapter-specific primer)	
sX12-5'_RACE1	ACCTTTCGACGAGGATG (5' RACE; reverse transcription)	*sX12*
sX12-5'_RACE2	CGACGAGGATGTGCAG (5' RACE; PCR)	*sX12*
sX12-3'_RACE	TGCACATCCTCGTCGA (3' RACE; PCR)	*sX12*

Oligonucleotide	Sequence (Purpose)	Gene
asX4-5'RACE1	GAATCTGGTCACGACG (5' RACE; reverse transcription)	*asX4*
asX4-5'RACE2	TCTGGTCACGACGGTG (5' RACE; PCR)	*asX4*
asX4-3'RACE	GTCGTGACCAGATTCCTTGTCT (3' RACE; PCR)	*asX4*
pCR2.1_colo_fw	ACGACGTTGTAAAACGACGG (colony PCR)	
pCR2.1_colo_rev	TTCACACAGGAAACAGCTATGAC (colony PCR)	
pCR2.1_seq	CACAGGAAACAGCTATGAC (sequencing)	
ΔsX12-left-fwd	CA<u>GGATCC</u>ATGATCAGGGTGTCGAGGTG (*sX12* deletion; left fragment)	*sX12*
ΔsX12-left-rev	AAC<u>AAGCTT</u>ATGACCTCTGCCGGTTGTCT (*sX12* deletion; left fragment)	*sX12*
ΔsX12-right-fwd	AAC<u>AAGCTT</u>AATCGATTTCCCGGCTTG (*sX12* deletion; right fragment)	*sX12*
ΔsX12-right-rev	GC<u>TCTAGA</u>GTGGCGACCTGGTACTTCAG (*sX12* deletion; right fragment)	*sX12*
sX12-comp-fwd	CA<u>GGATCC</u>AACGGAGTTCAGCACTACGGT (complementation of Δ*sX12*)	*sX12*
sX12-comp-rev	AAC<u>AAGCTT</u>CGAATTCTTGGAATGGTCA (complementation of Δ*sX12*)	*sX12*
sX6-fwd	TTT<u>GGTCT</u>CTATTCGCTGAGCCGATTTTTTGCAC (sX6-c-Myc tagging)	*sX6*
sX6-rev	TTT<u>GGTCT</u>CTCACCGCGTGCAGGCTTTTCTTCTTCC (sX6-c-Myc tagging)	*sX6*

[a], Ap, ampicillin; Gm, gentamycin; Km, kanamycin; Rif, rifampicin; Sm, spectinomycin; Tc, tetracycline. [R], resistance. Recognition sites of restriction enzymes are underlined.

*, RNA oligonucleotide carrying a phosphate modification at the 5' and an inverted deoxythymidine (idT) at the 3' end.

References

1. Canteros, B.I. (1990) Ph.D. thesis. *University of Florida, Gainesville, FL*.
2. Koebnik, R., Krüger, A., Thieme, F., Urban, A. and Bonas, U. (2006) Specific binding of the *Xanthomonas campestris* pv. vesicatoria AraC-type transcriptional activator HrpX to plant-inducible promoter boxes. *J. Bacteriol.*, **188**, 7652-7660.
3. Wengelnik, K., Rossier, O. and Bonas, U. (1999) Mutations in the regulatory gene *hrpG* of *Xanthomonas campestris* pv. vesicatoria result in constitutive expression of all *hrp* genes. *J. Bacteriol.*, **181**, 6828-6831.
4. Menard, R., Sansonetti, P.J. and Parsot, C. (1993) Nonpolar mutagenesis of the *ipa* genes defines IpaB, IpaC, and IpaD as effectors of *Shigella flexneri* entry into epithelial cells. *J. Bacteriol.*, **175**, 5899-5906.
5. Szczesny, R., Jordan, M., Schramm, C., Schulz, S., Cogez, V., Bonas, U. and Büttner, D. (2010) Functional characterization of the Xcs and Xps type II secretion systems from the plant pathogenic bacterium *Xanthomonas campestris* pv. *vesicatoria*. *New Phytol.*, **187**, 983-1002.
6. Bonas, U., Stall, R.E. and Staskawicz, B. (1989) Genetic and structural characterization of the avirulence gene *avrBs3* from *Xanthomonas campestris* pv. *vesicatoria*. *Mol. Gen. Genet.*, **218**, 127-136.
7. Huguet, E., Hahn, K., Wengelnik, K. and Bonas, U. (1998) *hpaA* mutants of *Xanthomonas campestris* pv. *vesicatoria* are affected in pathogenicity but retain the ability to induce host-specific hypersensitive reaction. *Mol. Microbiol.*, **29**, 1379-1390.
8. Figurski, D.H. and Helinski, D.R. (1979) Replication of an origin-containing derivative of plasmid RK2 dependent on a plasmid function provided *in trans*. *Proc. Natl. Acad. Sci. U. S. A.*, **76**, 1648-1652.

Table S2. Classification of putative TSSs which were automatically identified in the *Xcv* genome [see Figure 1B and METHODS; (1)].

TSS position[a]	Strand	p-value[b]	Library 2[c]	Library 1[c]	Primary[d] to CDS	5' UTR (bp)	Internal[d] to CDS	Distance to start/stop[e] (bp)	Antisense[d] to CDS	Distance to start/stop[e] (bp)	Orphan[f]
22092#	+	0.0001	6	0	*XCVd0018*	29	-	-	-	-	-
22614#	-	0.0034	7	2	-	-	-	-	*XCVd0019*	28/478	-
24098#	-	0.0191	3	0	-	-	-	-	-	-	+
31729#	-	0.0071	7	0	-	-	-	-	-	-	+
54664#	-	0.0344	3	0	-	-	-	-	*XCVd0049*	81/1463	-
56391#	-	0.0000	72	14	-	-	-	-	-	-	+
63600#	+	0.0003	6	0	-	-	-	-	*XCVd0054*	572/930	-
76432#	-	0.0007	7	0	*XCVd0069*	24	-	-	-	-	-
81111#	-	0.0032	9	2	-	-	-	-	*XCVd0073*	701/681	-
81749#	-	0.0014	9	2	-	-	-	-	*XCVd0073*	1339/43	-
82121#	+	0.0018	5	1	*XCVd0074*	21	-	-	-	-	-
88062#	+	0.0024	4	0	-	-	*XCVD0079*	757/2551	-	-	-
90670#	+	0.0341	9	0	*XCVd0080*	50	-	-	-	-	-
94989#	+	0.0081	4	0	-	-	*XCVD0086*	82/1066	-	-	-
95432#	-	0.0000	21	1	-	-	-	-	*XCVd0086*	525/623	-
97457#	-	0.0064	6	1	-	-	-	-	*XCVd0087*	1099/1000	-
99301#	-	0.0172	4	0	-	-	*XCVD0089*	238/370	-	-	-
100770#	+	0.0000	20	2	-	-	*XCVD0091*	817/13	-	-	-
101600#	+	0.0145	5	0	*XCVd0093*	287	-	-	*XCVd0092*	52/706	-
101864#	+	0.0000	156	25	*XCVd0093*	23	-	-	-	-	-
103085#	+	0.0369	4	3	-	-	-	-	*XCVd0094*	894/170	-
104001#	-	0.0000	57	11	*XCVd0094*	22	-	-	-	-	-
109851#	+	0.0000	35	6	-	-	-	-	*XCVd0099*	1283/99	-
112115#	-	0.0104	5	0	*XCVd0101* *XCVd0102*	285 28	-	-	-	-	-
112250#	+	0.0000	41	11	*XCVd0103*	27	-	-	-	-	-
113697#	+	0.0159	3	1	-	-	-	-	*XCVd0104*	755/579	-
114827#	-	0.0000	11	0	-	-	*XCVD0105*	786/281	-	-	-
115639#	-	0.0000	33	4	*XCVd0105*	26	-	-	-	-	-
116378#	+	0.0000	80	8	*XCVd0107*	277	-	-	*XCVd0106*	31/486	-
117900#	+	0.0000	18	2	-	-	-	-	*XCVd0109*	740/180	-
117969#	+	0.0018	13	2	-	-	-	-	*XCVd0109*	671/249	-
118654#	+	0.0156	8	1	*XCVd0110*	201	-	-	*XCVd0109*	14/934	-
118665#	-	0.0000	54	9	*XCVd0109*	25	-	-	-	-	-
119469#	+	0.0018	5	1	*XCVd0112*	49	*XCVd0111*	375/41	-	-	-
120806#	+	0.0083	6	1	-	-	*XCVd0112*	1288/1684	-	-	-
120867#	-	0.0229	5	1	-	-	-	-	*XCVd0112*	1349/1623	-
123136#	+	0.0004	5	0	*XCVd0114*	218	*XCVd0113*	624/71	-	-	-
127227#	-	0.0439	4	2	-	-	-	-	*XCVd0115*	3019/62	-
127749#	-	0.0000	28	3	-	-	-	-	*XCVd0116*	275/1032	-
136150#	+	0.0027	3	0	-	-	*XCVd0125*	512/210	-	-	-
137519#	-	0.0000	15	0	*XCVd0126*	23	-	-	-	-	-
138153#	+	0.0020	5	0	-	-	-	-	*XCVd0127*	74/355	-
151691#	+	0.0026	4	0	-	-	-	-	*XCVd0144*	337/304	-
159251#	+	0.0000	49	5	*XCVd0153*	36	-	-	*XCVd0152*	71/2254	-
159633#	+	0.0158	3	0	-	-	*XCVd0153*	346/49	*XCVd0154*	409/74	-
164247#	-	0.0168	3	0	-	-	*XCVd0155*	1062/4193	-	-	-
169348#	+	0.0024	3	0	-	-	-	-	*XCVd0160* *XCVd0161*	98/667 249/104	-
180140#	+	0.0049	3	0	-	-	-	-	*XCVd0171*	604/86	-
182433#	-	0.0050	3	0	*XCVd0172*	62	-	-	-	-	-

[a], TSS position on the *Xcv* chromosome and the plasmid (pXCV183; marked with #), respectively. [b], p-value indicates the significance of the annotated TSS (see METHODS). [c], number of read starts at TSS position. [d], TSS classification and the corresponding CDS. [e], distance of the TSS to annotated start and stop codon of the corresponding CDS. [f], a TSS is classified as orphan (+) if it is neither classified as primary, internal nor antisense TSS.

TSS position[a]	Strand	p-value[b]	Library 2[c]	Library 1[c]	Primary[d] to CDS	5' UTR (bp)	Internal[d] to CDS	Distance to start/stop[e] (bp)	Antisense[d] to CDS	Distance to start/stop[e] (bp)	Orphan[f]
1404	+	0.0000	15	1	XCV0002	202	-	-	-	-	-
9520	+	0.0078	5	2	XCV0007	68	-	-	-	-	-
10034	+	0.0036	4	0	-	-	XCV0007	446/744	-	-	-
10857	+	0.0001	10	1	XCV0008	78	-	-	-	-	-
11675	-	0.0001	6	0	-	-	-	-	XCV0009 XCV0008	17/775 740/72	-
15312	-	0.0000	12	0	XCV0013 XCV0014	297 20	-	-	-	-	-
17027	+	0.0000	42	18	-	-	XCV0016	23/900	-	-	-
17192	-	0.0001	12	0	-	-	-	-	XCV0016	188/735	-
18468	+	0.0022	3	0	-	-	XCV0017	538/49	-	-	-
18664	-	0.0147	3	0	-	-	-	-	-	-	+
23398	+	0.0000	13	1	-	-	XCV0022	45/317	-	-	-
38760	-	0.0000	13	0	-	-	XCV0035	12/425	-	-	-
44970	-	0.0000	14	3	XCV0037	79	-	-	-	-	-
45473	+	0.0246	5	0	XCV0039	58	-	-	XCV0038	1/376	-
45497	-	0.0170	3	0	XCV0038	23	-	-	XCV0039	34/1308	-
46397	-	0.0101	3	0	-	-	-	-	XCV0039	866/408	-
47391	-	0.0063	6	0	XCV0040	22	-	-	-	-	-
52187	+	0.0170	3	1	XCV0045	22	-	-	-	-	-
60944	-	0.0066	4	0	XCV0050	66	-	-	XCV0051	98/1191	-
62304	+	0.0206	3	0	XCV0052	23	-	-	-	-	-
72269	+	0.0213	3	0	XCV0061	116	-	-	XCV0060	27/761	-
72322	-	0.0002	9	1	XCV0060	80	-	-	XCV0061	63/404	-
72626	+	0.0001	13	2	-	-	XCV0061	241/100	-	-	-
75816	-	0.0016	4	0	-	-	-	-	-	-	+
77193	-	0.0006	6	1	XCV0064	16	-	-	-	-	-
78522	+	0.0000	23	2	-	-	XCV0066	817/13	XCV0067	231/46	-
78978	-	0.0187	8	5	XCV0067	225	-	-	-	-	-
81973	-	0.0123	6	2	XCV0069	270	-	-	-	-	-
84253	-	0.0047	3	0	XCV0072	20	-	-	-	-	-
84514	+	0.0058	3	0	-	-	XCV0073	121/460	-	-	-
85087	+	0.0000	11	0	-	-	-	-	-	-	+
87652	+	0.0469	7	0	XCV0077	22	-	-	-	-	-
89289	+	0.0065	4	0	-	-	XCV0078	1122/2840	-	-	-
103766	+	0.0206	3	0	-	-	-	-	XCV0088	747/149	-
109502	-	0.0009	6	1	XCV0092	72	-	-	-	-	-
117077	-	0.0026	4	0	XCV0099	274	XCV0100	573/266	-	-	-
118296	-	0.0401	3	1	-	-	XCV0101	675/524	-	-	-
133989	+	0.0164	3	1	-	-	XCV0117	8/588	-	-	-
138711	+	0.0011	6	1	XCV0120	26	-	-	-	-	-
139008	+	0.0333	3	0	-	-	XCV0120	271/934	-	-	-
149036	+	0.0002	7	0	-	-	-	-	-	-	+
149533	-	0.0000	13	1	-	-	XCV0125	817/13	-	-	-
152067	-	0.0152	5	0	-	-	-	-	XCV0128	155/543	-
155750	+	0.0005	11	2	XCV0131	227	-	-	-	-	-
156627	-	0.0000	15	3	-	-	-	-	XCV0132	106/295	-
192866	-	0.0118	3	0	XCV0160	31	-	-	-	-	-
193010	+	0.0008	9	2	XCV0161	0	-	-	-	-	-
193630	+	0.0149	3	1	-	-	XCV0161	620/750	-	-	-
195442	-	0.0000	22	0	XCV0162	23	-	-	-	-	-
196578	+	0.0000	37	10	XCV0164	32	-	-	-	-	-
201827	-	0.0015	7	2	XCV0168	0	-	-	XCV0169	12/1352	-
201899	+	0.0000	16	1	-	-	XCV0169	84/1280	XCV0168	72/1250	-

[a], TSS position on the *Xcv* chromosome and the plasmid (pXCV183; marked with #), respectively. [b], p-value indicates the significance of the annotated TSS (see METHODS). [c], number of read starts at TSS position. [d], TSS classification and the corresponding CDS. [e], distance of the TSS to annotated start and stop codon of the corresponding CDS. [f], a TSS is classified as orphan (+) if it is neither classified as primary, internal nor antisense TSS.

TSS position[a]	Strand	p-value[b]	Library 2[c]	Library 1[c]	Primary[d] to CDS	5' UTR (bp)	Internal[d] to CDS	Distance to start/stop[e] (bp)	Antisense[d] to CDS	Distance to start/stop[e] (bp)	Orphan[f]
204493	+	0.0000	12	2	-	-	-	-	XCV0172	3775/71	-
208848	+	0.0338	4	0	-	-	XCV0173	420/248	-	-	-
212756	+	0.0000	8	0	XCV0178	116	XCV0177	396/116	-	-	-
217013	+	0.0021	4	0	-	-	XCV0181	424/562	-	-	-
217960	+	0.0416	7	4	XCV0183	22	-	-	-	-	-
220222	+	0.0000	11	0	XCV0184	23	-	-	-	-	-
222838	-	0.0000	17	0	XCV0186	21	-	-	-	-	-
224042	+	0.0000	15	2	XCV0188	35	-	-	-	-	-
225642	+	0.0161	4	2	XCV0189	118	-	-	-	-	-
237714	-	0.0365	4	1	XCV0200	47	-	-	-	-	-
238550	-	0.0323	7	0	XCV0201	0	-	-	XCV0202	57/1010	-
240862	+	0.0050	6	2	XCV0205	23	-	-	-	-	-
241374	+	0.0011	4	0	XCV0206	55	-	-	-	-	-
244105	-	0.0000	194	14	XCV0208	62	-	-	-	-	-
244587	+	0.0000	15	2	-	-	XCV0209	4/1204	-	-	-
250048	+	0.0221	3	0	-	-	-	-	XCV0213	484/535	-
252996	+	0.0000	19	1	-	-	-	-	XCV0214	72/2471	-
253127	-	0.0000	25	10	XCV0214	203	-	-	-	-	-
261157	+	0.0000	17	2	XCV0221	81	-	-	-	-	-
261741	-	0.0000	15	1	-	-	XCV0222	817/13	XCV0221	503/3	-
265771	-	0.0000	43	8	-	-	XCV0226	216/569	-	-	-
266587	-	0.0000	66	8	XCV0227	90	-	-	-	-	-
267156	+	0.0000	17	2	-	-	-	-	-	-	+
269292	+	0.0043	3	0	-	-	-	-	-	-	+
278015	-	0.0012	6	0	XCV0236	0	-	-	-	-	-
278867	+	0.0100	5	0	XCV0238	23	-	-	-	-	-
280390	+	0.0019	3	0	XCV0240	22	-	-	-	-	-
283462	+	0.0037	4	0	XCV0243	0	-	-	-	-	-
285329	+	0.0140	4	1	XCV0244	22	-	-	-	-	-
287897	-	0.0105	4	1	XCV0246	21	-	-	-	-	-
288111	+	0.0128	10	5	XCV0247	0	-	-	-	-	-
292756	+	0.0016	8	0	XCV0252	122	-	-	XCV0251	108/629	-
295006	+	0.0046	6	1	XCV0255	119	-	-	-	-	-
304274	+	0.0276	5	0	XCV0264	150	-	-	XCV0263	49/904	-
307776	+	0.0002	11	2	XCV0266	37	-	-	-	-	-
311632	-	0.0178	3	0	-	-	XCV0268	98/213	-	-	-
315387	-	0.0003	5	0	XCV0270	36	-	-	-	-	-
320398	+	0.0021	6	1	XCV0274	234	-	-	XCV0273	105/1055	-
321787	-	0.0437	7	0	-	-	-	-	XCV0275 XCV0276	448/77 77/444	-
325673	+	0.0000	24	2	XCV0281	36	-	-	-	-	-
329286	-	0.0271	4	0	XCV0283	97	-	-	XCV0284	35/1287	-
331021	-	0.0263	6	1	-	-	XCV0285	933/278	-	-	-
331070	-	0.0003	10	1	-	-	XCV0285	884/327	-	-	-
332993	+	0.0000	9	1	XCV0287	0	-	-	XCV0286	67/939	-
337456	+	0.0001	13	3	XCV0293	22	-	-	-	-	-
338466	-	0.0057	5	2	-	-	XCV0294	940/259	-	-	-
341742	-	0.0051	5	2	XCV0296	23	-	-	-	-	-
341927	+	0.0000	12	0	XCV0297	26	-	-	-	-	-
344260	+	0.0000	27	4	XCV0299	26	-	-	-	-	-
347202	-	0.0014	4	0	-	-	XCV0300	99/2498	-	-	-
347842	-	0.0030	5	0	-	-	-	-	-	-	+
348352	+	0.0081	12	5	-	-	-	-	XCV0302	16/454	-
358119	-	0.0000	13	3	XCV0313	70	-	-	-	-	-

[a], TSS position on the *Xcv* chromosome and the plasmid (pXCV183; marked with #), respectively. [b], p-value indicates the significance of the annotated TSS (see METHODS). [c], number of read starts at TSS position. [d], TSS classification and the corresponding CDS. [e], distance of the TSS to annotated start and stop codon of the corresponding CDS. [f], a TSS is classified as orphan (+) if it is neither classified as primary, internal nor antisense TSS.

TSS position[a]	Strand	p-value[b]	Library 2[c]	Library 1[c]	Primary[d] to CDS	5' UTR (bp)	Internal[d] to CDS	Distance to start/stop[e] (bp)	Antisense[d] to CDS	Distance to start/stop[e] (bp)	Orphan[f]
369663	+	0.0028	3	0	-	-	XCV0322	30/833	-	-	-
375494	-	0.0081	3	0	-	-	XCV0327	24/425	XCV0328	90/1007	-
377721	+	0.0023	4	0	XCV0330	3	-	-	-	-	-
404292	-	0.0135	3	1	XCV0352	127	XCV0353	512/0	-	-	-
405223	-	0.0008	6	1	-	-	XCV0354	769/88	-	-	-
405567	+	0.0000	162	23	-	-	-	-	XCV0354	425/432	-
406317	-	0.0006	19	10	XCV0355	29	-	-	-	-	-
411941	+	0.0150	3	1	XCV0359	6	-	-	-	-	-
428338	-	0.0000	112	12	XCV0374	48	-	-	-	-	-
429323	-	0.0000	14	1	-	-	XCV0375	153/806	-	-	-
436669	+	0.0180	4	0	-	-	XCV0383	11/1368	-	-	-
441623	-	0.0130	4	0	-	-	XCV0387	26/879	-	-	-
441720	+	0.0000	11	0	XCV0388	30	-	-	XCV0387	71/976	-
447445	-	0.0001	6	2	XCV0392	52	-	-	-	-	-
448031	-	0.0000	12	2	XCV0393	23	-	-	-	-	-
448950	+	0.0098	3	0	XCV0395	226	XCV0394	784/172	-	-	-
449748	+	0.0205	5	1	XCV0396	23	-	-	-	-	-
453984	-	0.0000	24	7	-	-	XCV0400	168/758	-	-	-
455880	-	0.0012	5	1	XCV0402	24	-	-	-	-	-
458086	-	0.0001	6	0	-	-	-	-	-	-	+
462619	+	0.0014	4	0	-	-	-	-	XCV0411	2250/167	-
475337	-	0.0022	3	0	-	-	XCV0425	899/1020	-	-	-
477343	-	0.0002	8	0	XCV0426	25	-	-	-	-	-
477496	+	0.0000	22	3	XCV0427	37	-	-	-	-	-
485143	-	0.0000	13	0	-	-	-	-	XCV0435	1914/94	-
486264	-	0.0059	4	0	XCV0436	140	-	-	-	-	-
486511	+	0.0005	5	0	-	-	-	-	-	-	+
487914	-	0.0323	3	1	-	-	-	-	XCV0437	725/909	-
488081	+	0.0001	7	0	-	-	XCV0437	892/742	-	-	-
522402	-	0.0002	10	0	-	-	-	-	XCV0462	245/183	-
530343	+	0.0000	9	1	-	-	XCV0468	153/356	-	-	-
532431	-	0.0120	6	2	-	-	XCV0470	69/1289	-	-	-
532651	-	0.0025	5	0	XCV0470	151	-	-	-	-	-
532770	+	0.0009	8	1	XCV0471	0	-	-	-	-	-
533893	+	0.0033	6	0	-	-	-	-	XCV0472	302/36	-
537353	-	0.0066	4	0	XCV0475	32	XCV0476	1049/36	-	-	-
548017	+	0.0016	3	0	XCV0483	96	-	-	-	-	-
557964	+	0.0057	3	0	-	-	XCV0493	25/922	-	-	-
558968	-	0.0002	8	1	-	-	XCV0494	432/50	XCV0493	1029/82	-
568795	+	0.0038	5	1	XCV0504	27	-	-	-	-	-
573662	+	0.0040	3	0	XCV0509	53	XCV0508	156/50	-	-	-
581801	+	0.0005	4	0	XCV0514	94	-	-	-	-	-
585332	+	0.0311	8	5	XCV0518	0	XCV0517	790/4	-	-	-
586733	-	0.0003	8	0	-	-	XCV0519	327/362	-	-	-
587078	-	0.0000	62	8	XCV0519	18	-	-	-	-	-
587168	+	0.0000	39	0	XCV0520	30	-	-	-	-	-
587395	-	0.0439	3	0	-	-	-	-	XCV0520	197/594	-
589427	-	0.0000	10	0	-	-	XCV0522	42/650	-	-	-
589525	+	0.0000	73	23	XCV0523	85	-	-	XCV0522	56/748	-
590048	-	0.0012	10	0	-	-	-	-	XCV0523 XCV0524	438/13 7/382	-
590956	+	0.0000	24	1	XCV0525	24	-	-	-	-	-
592178	+	0.0000	18	5	XCV0527	24	-	-	-	-	-
598012	-	0.0331	4	2	XCV0531	22	-	-	XCV0532	86/853	-

[a], TSS position on the *Xcv* chromosome and the plasmid (pXCV183; marked with #), respectively. [b], p-value indicates the significance of the annotated TSS (see METHODS). [c], number of read starts at TSS position. [d], TSS classification and the corresponding CDS. [e], distance of the TSS to annotated start and stop codon of the corresponding CDS. [f], a TSS is classified as orphan (+) if it is neither classified as primary, internal nor antisense TSS.

TSS position[a]	Strand	p-value[b]	Library 2[c]	Library 1[c]	Primary[d] to CDS	5' UTR (bp)	Internal[d] to CDS	Distance to start/stop[e] (bp)	Antisense[d] to CDS	Distance to start/stop[e] (bp)	Orphan[f]
599405	+	0.0000	18	5	XCV0534	0	-	-	-	-	-
602321	+	0.0000	16	1	-	-	XCV0537	36/887	-	-	-
605417	-	0.0000	11	1	-	-	-	-	-	-	+
616013	-	0.0182	4	0	XCV0547	25	-	-	XCV0548	49/663	-
616018	-	0.0369	3	0	XCV0547	30	-	-	XCV0548	44/658	-
624601	-	0.0081	3	1	XCV0556	46	-	-	-	-	-
629187	+	0.0177	3	1	XCV0561	39	-	-	-	-	-
633751	-	0.0002	8	2	XCV0566	21	-	-	-	-	-
636699	-	0.0000	7	0	-	-	XCV0567	194/2121	-	-	-
647837	+	0.0000	18	0	XCV0574	192	-	-	-	-	-
654736	+	0.0074	4	0	-	-	-	-	-	-	+
656637	+	0.0079	6	2	-	-	XCV0579	42/2822	-	-	-
660302	+	0.0000	12	2	-	-	-	-	-	-	+
660601	-	0.0000	8	0	-	-	-	-	XCV0581	42/1880	-
668117	-	0.0002	6	0	XCV0586	52	-	-	XCV0587	83/670	-
687581	+	0.0000	24	2	-	-	XCV0608	268/1051	-	-	-
689806	-	0.0185	11	0	-	-	-	-	XCV0610	760/772	-
690527	+	0.0000	15	4	XCV0611	54	XCV0610	1481/51	-	-	-
692255	+	0.0469	3	0	-	-	XCV0611	1674/1010	-	-	-
693668	+	0.0006	4	0	XCV0612	26	-	-	-	-	-
694560	-	0.0169	3	0	-	-	-	-	XCV0612	866/111	-
698328	+	0.0000	14	2	XCV0616	34	-	-	-	-	-
703418	-	0.0086	3	0	-	-	-	-	-	-	+
707695	+	0.0343	4	1	-	-	-	-	XCV0625 XCV0624	838/8 79/258	-
707716	-	0.0000	13	1	XCV0624	100	XCV0625	817/13	-	-	-
713610	-	0.0193	3	0	-	-	XCV0632	9/1523	-	-	-
720541	-	0.0024	3	0	XCV0639	65	-	-	-	-	-
725332	+	0.0013	10	3	-	-	-	-	XCV0642	585/1994	-
732826	+	0.0424	4	0	-	-	XCV0648	43/334	-	-	-
736976	-	0.0027	3	0	-	-	XCV0653	8/789	-	-	-
738232	+	0.0005	12	4	XCV0656	25	-	-	-	-	-
739844	-	0.0122	3	0	-	-	XCV0657	862/121	-	-	-
740729	-	0.0000	12	2	XCV0657	23	-	-	-	-	-
749385	-	0.0005	6	1	-	-	XCV0665	168/956	-	-	-
752821	-	0.0000	16	1	XCV0666	23	-	-	-	-	-
758448	-	0.0056	5	0	XCV0669	34	-	-	-	-	-
758878	+	0.0469	7	0	-	-	-	-	XCV0670	1215/212	-
763156	+	0.0000	24	1	XCV0673	23	-	-	-	-	-
765474	+	0.0141	3	0	XCV0675	282	-	-	-	-	-
765583	+	0.0143	3	0	XCV0675	173	-	-	-	-	-
770365	+	0.0023	4	0	XCV0679	0	-	-	-	-	-
772840	+	0.0000	14	0	-	-	XCV0681	177/962	-	-	-
775297	-	0.0033	4	0	XCV0683	25	-	-	-	-	-
775992	+	0.0000	13	2	XCV0685	22	-	-	-	-	-
779535	-	0.0044	8	3	XCV0686	0	-	-	-	-	-
779651	+	0.0041	4	1	XCV0687	29	-	-	-	-	-
781259	-	0.0026	4	0	-	-	-	-	XCV0687	1579/490	-
782120	+	0.0002	7	0	XCV0688	0	-	-	-	-	-
782575	+	0.0000	13	2	XCV0689	244	-	-	-	-	-
785619	+	0.0000	18	2	XCV0692	17	-	-	-	-	-
793698	+	0.0000	11	1	XCV0700	23	-	-	-	-	-
799920	-	0.0039	6	2	-	-	XCV0705	769/88	XCV0704	942/49	-
800264	+	0.0000	162	23	-	-	-	-	XCV0705	425/432	-

[a], TSS position on the *Xcv* chromosome and the plasmid (pXCV183; marked with #), respectively. [b], p-value indicates the significance of the annotated TSS (see METHODS). [c], number of read starts at TSS position. [d], TSS classification and the corresponding CDS. [e], distance of the TSS to annotated start and stop codon of the corresponding CDS. [f], a TSS is classified as orphan (+) if it is neither classified as primary, internal nor antisense TSS.

TSS position[a]	Strand	p-value[b]	Library 2[c]	Library 1[c]	Primary[d] to CDS	5' UTR (bp)	Internal[d] to CDS	Distance to start/stop[e] (bp)	Antisense[d] to CDS	Distance to start/stop[e] (bp)	Orphan[f]
801014	-	0.0004	19	10	XCV0706	29	-	-	-	-	-
815890	+	0.0058	3	0	-	-	XCV0717	436/607	-	-	-
823961	+	0.0108	4	0	XCV0723	0	-	-	-	-	-
826922	+	0.0008	13	6	XCV0725	38	-	-	-	-	-
829714	+	0.0000	24	4	XCV0728	266	XCV0727	219/275	-	-	-
837085	-	0.0214	3	0	XCV0733	26	-	-	-	-	-
842923	+	0.0015	4	0	XCV0739	38	-	-	-	-	-
844848	+	0.0000	25	2	XCV0741	21	-	-	-	-	-
847069	+	0.0177	3	1	-	-	XCV0743	91/391	-	-	-
847628	+	0.0000	30	1	XCV0744	117	-	-	-	-	-
851592	+	0.0000	8	0	XCV0747	0	-	-	-	-	-
853751	+	0.0008	4	0	-	-	XCV0749	670/511	-	-	-
873841	+	0.0010	4	0	XCV0767	31	-	-	-	-	-
884284	-	0.0299	4	0	XCV0773	0	-	-	-	-	-
893392	+	0.0005	6	0	-	-	XCV0781	42/620	-	-	-
898261	-	0.0417	6	0	XCV0784	24	-	-	XCV0785	54/1151	-
900510	+	0.0380	4	0	-	-	XCV0787	97/292	-	-	-
905920	-	0.0007	5	0	XCV0792	60	-	-	-	-	-
906061	+	0.0000	16	3	XCV0793	24	-	-	-	-	-
906749	+	0.0000	10	0	XCV0794	31	-	-	-	-	-
907111	+	0.0010	5	0	-	-	XCV0794	331/919	-	-	-
907734	+	0.0007	6	0	-	-	XCV0794	954/296	-	-	-
931121	-	0.0030	3	0	-	-	XCV0816	41/324	-	-	-
933943	+	0.0000	86	27	-	-	-	-	-	-	+
935355	-	0.0000	11	1	XCV0819 XCV0820	264 16	-	-	-	-	-
936154	-	0.0000	12	1	-	-	XCV0821	15/764	-	-	-
936452	+	0.0048	11	1	XCV0822	18	-	-	-	-	-
936455	+	0.0028	23	4	XCV0822	15	-	-	-	-	-
946406	+	0.0206	3	0	-	-	XCV0830	571/721	-	-	-
951830	+	0.0001	10	2	XCV0835	181	-	-	-	-	-
953433	+	0.0005	12	1	XCV0836	51	-	-	-	-	-
956025	+	0.0000	14	0	XCV0839	101	-	-	-	-	-
956388	+	0.0055	5	0	-	-	XCV0839	262/2473	-	-	-
959187	+	0.0001	9	2	XCV0840	0	-	-	-	-	-
963192	-	0.0156	3	1	XCV0843	22	-	-	-	-	-
963547	+	0.0026	4	0	XCV0844	20	-	-	-	-	-
964753	+	0.0000	15	1	XCV0845	31	-	-	-	-	-
973152	+	0.0000	80	8	-	-	-	-	-	-	+
973732	+	0.0349	3	0	-	-	XCV0854	279/650	-	-	-
976232	+	0.0206	3	0	-	-	-	-	XCV0856	647/792	-
977026	-	0.0014	5	0	XCV0856	147	-	-	-	-	-
994684	-	0.0007	4	0	XCV0865	33	-	-	-	-	-
997844	-	0.0000	21	0	XCV0868	0	-	-	-	-	-
1003187	-	0.0005	5	0	XCV0871	27	-	-	-	-	-
1017977	+	0.0000	67	8	XCV0887	22	-	-	-	-	-
1021354	+	0.0006	11	3	XCV0890	73	-	-	-	-	-
1024293	+	0.0000	11	2	XCV0894	27	-	-	-	-	-
1028253	-	0.0039	3	0	XCV0899	0	-	-	-	-	-
1029078	-	0.0008	5	0	XCV0900	0	-	-	-	-	-
1031658	-	0.0108	3	1	XCV0902	80	XCV0903	2346/80	-	-	-
1034027	-	0.0212	4	2	XCV0903	23	-	-	-	-	-
1036242	-	0.0129	5	0	XCV0906	23	-	-	-	-	-
1036560	+	0.0006	6	0	XCV0907	0	-	-	-	-	-

[a], TSS position on the *Xcv* chromosome and the plasmid (pXCV183; marked with #), respectively. [b], p-value indicates the significance of the annotated TSS (see METHODS). [c], number of read starts at TSS position. [d], TSS classification and the corresponding CDS. [e], distance of the TSS to annotated start and stop codon of the corresponding CDS. [f], a TSS is classified as orphan (+) if it is neither classified as primary, internal nor antisense TSS.

TSS position[a]	Strand	p-value[b]	Library 2[c]	Library 1[c]	Primary[d] to CDS	5' UTR (bp)	Internal[d] to CDS	Distance to start/stop[e] (bp)	Antisense[d] to CDS	Distance to start/stop[e] (bp)	Orphan[f]
1041488	-	0.0166	3	1	-	-	-	-	XCV0911	1007/21	-
1042407	+	0.0004	7	1	XCV0913	180	-	-	XCV0912	54/725	-
1050250	-	0.0003	6	0	-	-	XCV0920	152/429	-	-	-
1052158	-	0.0193	3	0	XCV0922	21	-	-	-	-	-
1052229	+	0.0095	4	1	XCV0923	32	-	-	XCV0922	92/994	-
1055953	-	0.0005	5	0	XCV0926	26	-	-	-	-	-
1059633	-	0.0003	12	3	XCV0930	23	-	-	-	-	-
1064780	+	0.0131	3	1	XCV0937	26	-	-	-	-	-
1067662	-	0.0119	3	0	-	-	XCV0939	104/861	-	-	-
1073410	+	0.0000	17	0	XCV0945	19	-	-	-	-	-
1086476	+	0.0025	4	0	-	-	XCV0957	74/585	-	-	-
1095646	-	0.0012	4	0	-	-	-	-	XCV0965 XCV0964	12/1205 334/86	-
1102907	-	0.0272	3	1	XCV0973	0	-	-	-	-	-
1107405	+	0.0024	4	0	XCV0977	35	-	-	XCV0976	31/1326	-
1107410	-	0.0000	8	0	XCV0976	36	-	-	XCV0977	30/1715	-
1110383	+	0.0004	16	5	-	-	XCV0979	623/261	-	-	-
1114497	+	0.0001	8	0	-	-	-	-	-	-	+
1116481	+	0.0001	7	2	XCV0986	259	XCV0985	162/242	-	-	-
1117349	+	0.0000	9	1	XCV0987	147	-	-	-	-	-
1117882	+	0.0012	7	2	XCV0988	46	XCV0987	386/39	-	-	-
1118456	+	0.0303	5	0	-	-	XCV0988	528/167	-	-	-
1118678	+	0.0000	46	10	-	-	-	-	-	-	+
1120586	+	0.0007	4	1	-	-	XCV0991	172/3976	-	-	-
1122235	+	0.0042	3	1	-	-	XCV0991	1821/2327	-	-	-
1122779	+	0.0000	26	1	-	-	XCV0991	2365/1783	-	-	-
1128623	+	0.0074	5	1	-	-	XCV0992	3957/254	-	-	-
1129024	+	0.0000	10	1	XCV0993	75	-	-	-	-	-
1130406	+	0.0005	5	0	-	-	XCV0995	315/1799	-	-	-
1134713	-	0.0133	3	0	-	-	-	-	-	-	+
1135113	+	0.0030	15	12	-	-	-	-	-	-	+
1137159	+	0.0000	79	14	XCV1002	180	XCV1001	130/166	-	-	-
1137964	+	0.0008	9	2	XCV1003	209	XCV1002	625/199	-	-	-
1142119	+	0.0000	38	13	XCV1013	108	-	-	-	-	-
1146342	+	0.0000	37	12	-	-	-	-	-	-	+
1149059	+	0.0004	4	1	XCV1024	233	XCV1023	946/49	-	-	-
1149619	+	0.0026	3	0	XCV1025	180	XCV1024	327/53	-	-	-
1150093	+	0.0000	44	3	XCV1026	23	XCV1025	294/17	-	-	-
1153493	+	0.0000	10	1	XCV1028	17	-	-	-	-	-
1160474	-	0.0000	13	2	XCV1032	43	-	-	-	-	-
1160640	+	0.0000	20	1	XCV1033	33	-	-	-	-	-
1162992	-	0.0000	35	12	XCV1035	22	-	-	-	-	-
1164003	-	0.0006	5	0	XCV1036	163	-	-	-	-	-
1168793	+	0.0000	17	8	XCV1042	80	-	-	-	-	-
1169205	+	0.0000	19	3	-	-	XCV1042	332/762	-	-	-
1170131	+	0.0000	16	0	XCV1043	32	-	-	-	-	-
1173134	-	0.0006	5	0	XCV1046	22	-	-	-	-	-
1177286	+	0.0126	3	0	XCV1051	145	-	-	-	-	-
1179872	-	0.0114	5	2	XCV1052	40	-	-	-	-	-
1184957	+	0.0000	12	1	XCV1055	25	-	-	-	-	-
1185123	-	0.0387	7	2	-	-	-	-	XCV1055	141/491	-
1189664	+	0.0048	5	0	XCV1060	99	-	-	-	-	-
1190312	+	0.0006	8	2	XCV1061	228	XCV1060	549/194	-	-	-
1206581	-	0.0144	3	0	-	-	XCV1073	42/1661	XCV1074	99/437	-

[a], TSS position on the *Xcv* chromosome and the plasmid (pXCV183; marked with #), respectively. [b], p-value indicates the significance of the annotated TSS (see METHODS). [c], number of read starts at TSS position. [d], TSS classification and the corresponding CDS. [e], distance of the TSS to annotated start and stop codon of the corresponding CDS. [f], a TSS is classified as orphan (+) if it is neither classified as primary, internal nor antisense TSS.

TSS position[a]	Strand	p-value[b]	Library 2[c]	Library 1[c]	Primary[d] to CDS	5' UTR (bp)	Internal[d] to CDS	Distance to start/stop[e] (bp)	Antisense[d] to CDS	Distance to start/stop[e] (bp)	Orphan[f]
1206687	+	0.0000	11	1	-	-	XCV1074	7/331	XCV1073	64/1767	-
1207568	+	0.0000	8	0	XCV1076	72	-	-	-	-	-
1209562	-	0.0027	5	0	XCV1077	20	-	-	-	-	-
1209662	+	0.0017	10	0	-	-	-	-	-	-	+
1217572	+	0.0134	3	1	-	-	XCV1083	1641/827	-	-	-
1218498	+	0.0000	46	8	XCV1084	118	-	-	-	-	-
1219137	-	0.0112	3	0	-	-	-	-	-	-	+
1219503	+	0.0007	6	1	XCV1085	25	-	-	-	-	-
1223677	-	0.0046	3	0	-	-	XCV1087	426/338	-	-	-
1226768	-	0.0027	5	1	XCV1091	27	-	-	-	-	-
1227290	+	0.0000	10	0	-	-	XCV1092	172/1306	-	-	-
1233373	+	0.0000	19	2	-	-	XCV1098	817/13	-	-	-
1233578	+	0.0000	28	5	-	-	-	-	-	-	+
1235373	+	0.0020	5	1	-	-	-	-	-	-	+
1240320	+	0.0038	6	1	XCV1107	20	-	-	-	-	-
1242175	-	0.0000	91	11	-	-	-	-	-	-	+
1244521	-	0.0000	16	1	XCV1110	22	-	-	-	-	-
1249118	+	0.0002	20	9	XCV1118	29	XCV1117	910/16	-	-	-
1249868	-	0.0000	170	21	-	-	-	-	XCV1119	425/432	-
1250212	+	0.0041	5	1	XCV1120	153	XCV1119	769/88	-	-	-
1257067	-	0.0053	3	0	-	-	XCV1126	180/1334	-	-	-
1257541	+	0.0001	14	4	XCV1127	21	-	-	-	-	-
1261798	-	0.0002	16	6	XCV1131	65	XCV1132	1884/65	-	-	-
1265655	-	0.0171	3	1	XCV1134	0	-	-	XCV1135	40/1807	-
1270936	+	0.0000	47	4	-	-	XCV1141	37/517	-	-	-
1272952	-	0.0053	3	0	-	-	-	-	-	-	+
1273770	+	0.0000	36	10	XCV1145	23	-	-	-	-	-
1278951	+	0.0001	8	0	XCV1150	29	-	-	-	-	-
1284606	-	0.0088	4	1	-	-	XCV1155	27/1619	-	-	-
1287704	+	0.0032	3	0	-	-	XCV1158	694/1897	-	-	-
1295774	+	0.0131	4	1	XCV1164	42	XCV1163	308/12	-	-	-
1298129	-	0.0224	3	0	-	-	-	-	XCV1164	2313/2585	-
1309475	+	0.0005	8	1	XCV1172	65	-	-	XCV1171	98/196	-
1312832	+	0.0002	9	0	XCV1175	61	-	-	-	-	-
1321929	-	0.0414	5	0	XCV1181	22	-	-	XCV1182	50/637	-
1340207	+	0.0008	6	1	XCV1198	233	XCV1197	282/212	-	-	-
1343656	+	0.0005	5	0	XCV1202	63	-	-	-	-	-
1364483	+	0.0191	5	2	XCV1217	20	-	-	-	-	-
1368267	-	0.0345	6	3	XCV1221 XCV1222	249 19	-	-	-	-	-
1376702	+	0.0016	8	2	-	-	XCV1228	90/857	-	-	-
1381773	+	0.0121	3	0	XCV1232	23	-	-	-	-	-
1384583	-	0.0028	7	2	XCV1233	20	-	-	-	-	-
1384820	+	0.0000	17	3	XCV1234	46	-	-	-	-	-
1386923	+	0.0068	3	0	-	-	XCV1236	66/1907	-	-	-
1390988	-	0.0193	3	0	-	-	XCV1239	35/150	-	-	-
1399557	-	0.0435	4	2	XCV1243	291	-	-	-	-	-
1404809	+	0.0275	3	0	XCV1247	89	-	-	-	-	-
1411639	-	0.0000	18	4	XCV1251	22	-	-	-	-	-
1419087	+	0.0005	8	1	XCV1260	0	-	-	XCV1259	78/833	-
1421227	+	0.0000	11	0	XCV1262	48	-	-	-	-	-
1424474	-	0.0060	4	0	XCV1265	171	-	-	-	-	-
1426111	-	0.0236	10	0	XCV1268	21	-	-	-	-	-
1429267	-	0.0016	4	0	XCV1271	24	-	-	-	-	-

[a], TSS position on the *Xcv* chromosome and the plasmid (pXCV183; marked with #), respectively. [b], p-value indicates the significance of the annotated TSS (see METHODS). [c], number of read starts at TSS position. [d], TSS classification and the corresponding CDS. [e], distance of the TSS to annotated start and stop codon of the corresponding CDS. [f], a TSS is classified as orphan (+) if it is neither classified as primary, internal nor antisense TSS.

TSS position[a]	Strand	p-value[b]	Library 2[c]	Library 1[c]	Primary[d] to CDS	5' UTR (bp)	Internal[d] to CDS	Distance to start/stop[e] (bp)	Antisense[d] to CDS	Distance to start/stop[e] (bp)	Orphan[f]
1433010	-	0.0000	18	5	XCV1275	29	-	-	-	-	-
1436219	-	0.0000	8	0	XCV1279	24	-	-	-	-	-
1437031	+	0.0023	6	0	XCV1281	28	-	-	XCV1280	4/763	-
1437682	-	0.0023	3	0	-	-	XCV1282	365/36	-	-	-
1440046	-	0.0054	3	0	-	-	XCV1283	964/1996	-	-	-
1441077	-	0.0028	5	0	XCV1283	67	-	-	-	-	-
1443612	-	0.0000	49	7	XCV1287	41	-	-	-	-	-
1443614	-	0.0007	48	7	XCV1287	43	-	-	-	-	-
1445416	+	0.0000	17	4	XCV1289	51	-	-	-	-	-
1450021	+	0.0001	8	1	XCV1292	0	-	-	-	-	-
1451420	+	0.0001	14	1	-	-	-	-	-	-	+
1451553	+	0.0055	10	2	-	-	-	-	-	-	+
1454260	-	0.0125	3	1	-	-	-	-	XCV1295	821/15	-
1456955	+	0.0000	37	0	XCV1299	25	-	-	-	-	-
1457771	+	0.0000	12	2	-	-	XCV1299	791/84	-	-	-
1458095	-	0.0000	15	1	-	-	XCV1300	819/11	-	-	-
1460678	+	0.0000	9	0	-	-	-	-	XCV1303	554/672	-
1461842	-	0.0113	3	1	XCV1304	23	-	-	-	-	-
1463947	-	0.0002	7	1	XCV1306	112	-	-	-	-	-
1464022	+	0.0151	9	0	-	-	-	-	-	-	+
1464279	+	0.0000	107	6	XCV1307	118	-	-	-	-	-
1464610	+	0.0122	3	0	-	-	XCV1307	213/719	-	-	-
1468435	+	0.0000	36	10	XCV1311	30	-	-	-	-	-
1470573	+	0.0005	6	0	XCV1312	22	-	-	-	-	-
1498473	-	0.0002	8	0	-	-	XCV1332	8/3195	-	-	-
1498586	+	0.0000	18	3	XCV1333	21	-	-	-	-	-
1502160	+	0.0000	15	3	XCV1336	21	-	-	-	-	-
1504286	-	0.0000	15	0	XCV1337	0	-	-	-	-	-
1504669	+	0.0019	5	1	XCV1338	28	-	-	-	-	-
1506240	+	0.0023	7	2	XCV1339	21	-	-	-	-	-
1508171	+	0.0000	62	7	XCV1342	45	-	-	-	-	-
1508253	-	0.0275	8	0	-	-	-	-	XCV1342	37/217	-
1508464	+	0.0113	6	1	XCV1343	23	XCV1342	248/6	-	-	-
1509902	+	0.0000	34	3	XCV1345	97	-	-	-	-	-
1529221	+	0.0200	8	0	-	-	XCV1357	388/1945	-	-	-
1543610	+	0.0094	6	0	-	-	XCV1365	301/493	-	-	-
1546532	+	0.0000	41	11	XCV1368	22	-	-	-	-	-
1549846	-	0.0000	19	0	XCV1369	30	-	-	-	-	-
1549864	+	0.0081	6	1	XCV1370	189	-	-	XCV1369	48/2126	-
1550020	+	0.0162	5	1	XCV1370	33	-	-	-	-	-
1553394	+	0.0000	25	3	XCV1373	50	-	-	-	-	-
1555991	+	0.0086	7	2	XCV1375	180	XCV1374	651/146	-	-	-
1567734	+	0.0048	3	0	XCV1387	38	XCV1386	1047/35	-	-	-
1573503	-	0.0021	3	0	-	-	XCV1393	482/2016	-	-	-
1580759	-	0.0050	6	1	XCV1399	248	XCV1400	322/16	-	-	-
1581048	-	0.0441	5	1	-	-	XCV1400	33/305	-	-	-
1582399	-	0.0000	14	2	XCV1401	20	-	-	-	-	-
1582406	+	0.0046	5	1	XCV1402	128	-	-	XCV1401	27/1199	-
1589016	+	0.0002	9	1	XCV1406	69	-	-	-	-	-
1589822	+	0.0000	10	1	XCV1408	22	-	-	-	-	-
1590522	-	0.0020	3	0	-	-	XCV1409	338/225	-	-	-
1591652	-	0.0054	5	1	XCV1410	154	-	-	-	-	-
1593751	-	0.0000	7	0	XCV1413	68	-	-	-	-	-
1595169	+	0.0009	7	0	XCV1415	21	-	-	XCV1414	76/1140	-

[a], TSS position on the *Xcv* chromosome and the plasmid (pXCV183; marked with #), respectively. [b], p-value indicates the significance of the annotated TSS (see METHODS). [c], number of read starts at TSS position. [d], TSS classification and the corresponding CDS. [e], distance of the TSS to annotated start and stop codon of the corresponding CDS. [f], a TSS is classified as orphan (+) if it is neither classified as primary, internal nor antisense TSS.

TSS position[a]	Strand	p-value[b]	Library 2[c]	Library 1[c]	Primary[d] to CDS	5' UTR (bp)	Internal[d] to CDS	Distance to start/stop[e] (bp)	Antisense[d] to CDS	Distance to start/stop[e] (bp)	Orphan[f]
1601206	+	0.0360	5	2	XCV1423 XCV1424	0 264	-	-	-	-	-
1601450	+	0.0022	5	0	XCV1424	20	-	-	-	-	-
1609450	+	0.0000	17	0	-	-	-	-	XCV1430	151/391	-
1614026	-	0.0069	3	0	-	-	XCV1434	128/1002	-	-	-
1618771	-	0.0010	5	1	XCV1440	52	-	-	-	-	-
1618915	+	0.0038	4	0	-	-	-	-	XCV1441	2116/35	-
1624149	-	0.0192	5	0	-	-	-	-	XCV1446	1/678	-
1629856	-	0.0009	5	0	XCV1450	22	-	-	-	-	-
1631636	-	0.0147	3	0	XCV1452	8	-	-	-	-	-
1631992	+	0.0137	3	1	XCV1453	10	-	-	-	-	-
1639397	+	0.0060	4	1	XCV1459	28	-	-	XCV1458	46/645	-
1641414	-	0.0122	3	0	XCV1461	44	-	-	-	-	-
1642409	-	0.0000	20	5	XCV1462	22	-	-	-	-	-
1644256	-	0.0088	3	0	-	-	XCV1463	1905/1682	-	-	-
1648431	-	0.0007	5	0	-	-	XCV1465	8/1308	-	-	-
1653692	-	0.0000	18	6	XCV1470	209	XCV1471	1221/122	-	-	-
1658315	-	0.0000	49	5	XCV1475	25	-	-	-	-	-
1659232	-	0.0166	3	1	XCV1476	32	-	-	-	-	-
1659403	+	0.0034	4	0	-	-	-	-	XCV1477	1427/93	-
1660847	-	0.0038	6	2	XCV1477	17	-	-	-	-	-
1663194	-	0.0000	15	4	XCV1479	242	-	-	-	-	-
1669473	+	0.0000	12	1	XCV1485	0	-	-	-	-	-
1676072	+	0.0012	7	0	XCV1490	220	-	-	-	-	-
1684643	+	0.0000	70	11	XCV1495	21	-	-	-	-	-
1690056	+	0.0113	4	0	XCV1499	63	-	-	XCV1498	51/899	-
1690128	+	0.0290	3	0	-	-	XCV1499	9/437	-	-	-
1700939	-	0.0019	3	0	XCV1506	26	-	-	XCV1507	39/1895	-
1702829	+	0.0009	5	0	XCV1508	16	XCV1507	1929/5	-	-	-
1704019	-	0.0037	4	0	XCV1509	69	-	-	-	-	-
1704140	+	0.0005	8	1	XCV1510	38	-	-	-	-	-
1708364	-	0.0001	18	7	XCV1513	70	-	-	-	-	-
1710153	-	0.0031	3	0	XCV1515	25	-	-	-	-	-
1712207	-	0.0034	4	0	XCV1516	0	-	-	-	-	-
1712476	+	0.0277	4	2	XCV1517	30	-	-	-	-	-
1719682	+	0.0000	20	2	XCV1525	0	-	-	XCV1524	75/500	-
1721663	+	0.0000	20	5	XCV1528	34	-	-	-	-	-
1724343	-	0.0193	3	0	-	-	-	-	XCV1530	504/1025	-
1726530	+	0.0027	3	0	XCV1532	29	-	-	-	-	-
1733457	+	0.0000	12	2	XCV1537	39	-	-	-	-	-
1736086	+	0.0264	3	0	XCV1539	131	-	-	XCV1538	1/1030	-
1745530	+	0.0024	3	0	-	-	XCV1545	189/41	-	-	-
1750856	+	0.0000	21	2	-	-	XCV1551	817/13	XCV1552	537/46	-
1753470	-	0.0000	31	1	-	-	-	-	-	-	+
1756674	+	0.0000	18	6	XCV1557	73	-	-	XCV1556	12/488	-
1756686	-	0.0000	27	2	XCV1556	0	-	-	XCV1557	61/489	-
1758176	-	0.0000	17	1	XCV1559	24	-	-	XCV1560	81/488	-
1758200	+	0.0493	3	0	XCV1560	57	-	-	XCV1559	48/440	-
1762031	+	0.0060	8	4	XCV1563	36	-	-	-	-	-
1762679	+	0.0002	7	1	XCV1564	48	-	-	-	-	-
1768866	-	0.0152	5	0	-	-	XCV1568	81/197	-	-	-
1778677	-	0.0061	3	0	-	-	XCV1576	903/536	-	-	-
1779063	-	0.0195	3	0	-	-	XCV1576	517/922	-	-	-
1783856	-	0.0000	18	5	-	-	XCV1578	65/2904	-	-	-

[a], TSS position on the *Xcv* chromosome and the plasmid (pXCV183; marked with #), respectively. [b], p-value indicates the significance of the annotated TSS (see METHODS). [c], number of read starts at TSS position. [d], TSS classification and the corresponding CDS. [e], distance of the TSS to annotated start and stop codon of the corresponding CDS. [f], a TSS is classified as orphan (+) if it is neither classified as primary, internal nor antisense TSS.

TSS position[a]	Strand	p-value[b]	Library 2[c]	Library 1[c]	Primary[d] to CDS	5' UTR (bp)	Internal[d] to CDS	Distance to start/stop[e] (bp)	Antisense[d] to CDS	Distance to start/stop[e] (bp)	Orphan[f]
1783909	+	0.0004	5	0	-	-	-	-	XCV1578 XCV1579	12/2957 978/52	-
1788756	-	0.0008	4	0	XCV1582	28	-	-	XCV1583	7/1117	-
1791356	-	0.0000	13	0	XCV1584	147	-	-	-	-	-
1793470	-	0.0004	7	0	-	-	XCV1586	5/441	-	-	-
1797706	-	0.0000	9	0	XCV1591	280	-	-	-	-	-
1798406	+	0.0000	19	7	XCV1593	47	-	-	-	-	-
1803085	+	0.0019	5	1	XCV1598	11	-	-	-	-	-
1814647	+	0.0433	9	0	-	-	XCV1606	149/1542	-	-	-
1816736	+	0.0000	17	0	XCV1607	1	-	-	-	-	-
1819516	+	0.0183	3	0	XCV1608	23	-	-	-	-	-
1820299	+	0.0000	12	0	XCV1609	22	-	-	-	-	-
1821313	+	0.0001	9	1	-	-	XCV1610	140/3054	-	-	-
1825238	-	0.0001	8	1	-	-	XCV1611	6/632	-	-	-
1830728	-	0.0186	3	0	-	-	XCV1617	37/1051	-	-	-
1833891	+	0.0000	87	4	-	-	XCV1620	63/338	-	-	-
1838797	+	0.0000	40	12	XCV1626	51	-	-	-	-	-
1843191	+	0.0014	5	0	XCV1631	0	-	-	XCV1630	24/596	-
1859472	-	0.0028	7	2	XCV1646	0	-	-	-	-	-
1859602	+	0.0008	6	1	XCV1647	26	-	-	-	-	-
1864572	+	0.0002	7	0	-	-	XCV1654	8/699	-	-	-
1868556	-	0.0000	40	4	XCV1659	30	-	-	-	-	-
1868712	+	0.0002	7	1	XCV1660	19	-	-	-	-	-
1874518	+	0.0001	15	4	XCV1665	38	-	-	-	-	-
1875752	-	0.0005	6	0	-	-	-	-	XCV1666	458/78	-
1880820	-	0.0193	3	0	-	-	-	-	XCV1669 XCV1670	1001/42 70/753	-
1882951	+	0.0000	12	3	XCV1672	98	-	-	-	-	-
1895289	-	0.0047	5	1	XCV1679	25	-	-	XCV1680	36/1403	-
1895299	+	0.0072	4	0	XCV1680	26	-	-	XCV1679	35/1237	-
1900120	-	0.0065	5	1	XCV1684	90	-	-	-	-	-
1903361	+	0.0000	26	5	XCV1689	21	-	-	-	-	-
1908861	-	0.0008	5	0	XCV1693	19	-	-	XCV1694	93/1370	-
1908930	+	0.0000	31	1	XCV1694	24	-	-	XCV1693	88/756	-
1911229	+	0.0005	7	1	XCV1696	12	-	-	XCV1695	89/670	-
1912983	+	0.0148	3	1	-	-	XCV1697	180/1178	-	-	-
1914145	+	0.0003	19	9	XCV1698	113	XCV1697	1342/16	-	-	-
1914895	-	0.0000	170	21	-	-	-	-	XCV1699	425/432	-
1915239	+	0.0039	5	1	-	-	XCV1699	769/88	-	-	-
1921894	-	0.0000	44	2	-	-	XCV1704	297/491	-	-	-
1925949	+	0.0000	11	1	XCV1707	18	-	-	-	-	-
1938634	-	0.0204	3	1	XCV1719	95	-	-	-	-	-
1961959	+	0.0022	3	0	XCV1741	21	-	-	-	-	-
1964107	-	0.0151	3	0	-	-	XCV1743	400/937	-	-	-
1964739	+	0.0021	3	0	XCV1744	41	-	-	-	-	-
1966625	+	0.0056	3	0	XCV1746	28	-	-	-	-	-
1971387	-	0.0000	8	0	-	-	XCV1748	108/1910	-	-	-
1971505	+	0.0000	15	1	-	-	-	-	XCV1748	10/2028	-
1972062	+	0.0000	13	1	XCV1749	22	-	-	-	-	-
1974295	-	0.0437	7	0	-	-	-	-	XCV1750	326/501	-
1975236	+	0.0034	4	0	XCV1752	47	-	-	-	-	-
1976107	+	0.0055	5	0	-	-	XCV1752	824/465	-	-	-
1980337	+	0.0000	22	0	XCV1758	17	-	-	XCV1757	83/622	-
1984126	-	0.0005	4	0	XCV1763	24	-	-	XCV1764	42/671	-

[a], TSS position on the *Xcv* chromosome and the plasmid (pXCV183; marked with #), respectively. [b], p-value indicates the significance of the annotated TSS (see METHODS). [c], number of read starts at TSS position. [d], TSS classification and the corresponding CDS. [e], distance of the TSS to annotated start and stop codon of the corresponding CDS. [f], a TSS is classified as orphan (+) if it is neither classified as primary, internal nor antisense TSS.

TSS position[a]	Strand	p-value[b]	Library 2[c]	Library 1[c]	Primary[d] to CDS	5' UTR (bp)	Internal[d] to CDS	Distance to start/stop[e] (bp)	Antisense[d] to CDS	Distance to start/stop[e] (bp)	Orphan[f]
1984810	+	0.0051	8	3	XCV1765	45	-	-	-	-	-
1988882	+	0.0042	9	2	XCV1769 XCV1768	2999	-	-	-	-	-
1989065	+	0.0275	4	0	XCV1769	116	XCV1768	174/101	-	-	-
1993650	+	0.0050	4	1	XCV1772	0	-	-	-	-	-
1996311	+	0.0230	7	3	XCV1775	74	-	-	-	-	-
1996655	+	0.0419	5	1	-	-	XCV1775	270/2375	-	-	-
1998924	+	0.0002	18	1	XCV1776	249	XCV1775	2539/106	-	-	-
1999090	+	0.0000	30	4	XCV1776	83	-	-	-	-	-
2011844	-	0.0138	4	0	XCV1789	264	XCV1790	374/249	-	-	-
2012054	-	0.0424	4	0	-	-	XCV1790	164/459	-	-	-
2014218	-	0.0000	15	3	XCV1793	19	-	-	-	-	-
2018225	+	0.0012	6	1	XCV1798	30	-	-	-	-	-
2041528	+	0.0004	8	2	XCV1810	30	-	-	-	-	-
2046892	-	0.0002	8	1	-	-	-	-	XCV1814	28/1392	-
2046910	+	0.0095	3	0	XCV1814	10	-	-	-	-	-
2048788	+	0.0014	5	1	XCV1816	24	-	-	-	-	-
2050617	+	0.0000	11	0	XCV1818	26	-	-	-	-	-
2056210	+	0.0105	3	0	XCV1823	3	-	-	-	-	-
2061162	+	0.0026	5	1	-	-	-	-	XCV1825	1738/259	-
2063383	+	0.0003	5	0	-	-	XCV1826	165/809	-	-	-
2064776	-	0.0056	3	0	-	-	XCV1827	166/397	-	-	-
2066165	-	0.0265	6	3	XCV1828	0	-	-	-	-	-
2066489	+	0.0029	7	0	XCV1829	0	-	-	-	-	-
2076935	-	0.0010	4	0	XCV1838	55	-	-	-	-	-
2077009	+	0.0344	6	0	-	-	-	-	-	-	+
2077238	+	0.0116	7	1	XCV1839	103	-	-	-	-	-
2078061	+	0.0225	4	1	-	-	-	-	XCV1840	838/8	-
2078080	-	0.0000	14	1	-	-	XCV1840	819/11	-	-	-
2080165	+	0.0001	20	9	XCV1843	29	-	-	XCV1842	17/930	-
2080915	-	0.0000	170	21	-	-	-	-	XCV1844	425/432	-
2081259	+	0.0038	5	1	XCV1845	156	XCV1844	769/88	-	-	-
2081955	+	0.0103	3	0	-	-	XCV1845	540/11	XCV1846	816/76	-
2088259	+	0.0013	5	0	XCV1852	36	-	-	-	-	-
2094165	+	0.0004	6	0	XCV1859	103	-	-	-	-	-
2100728	+	0.0004	5	0	-	-	XCV1861	3048/8078	-	-	-
2108278	-	0.0259	3	0	-	-	-	-	XCV1861	10598/528	-
2109973	+	0.0000	20	2	XCV1864	151	XCV1863	817/13	-	-	-
2111307	+	0.0000	17	4	XCV1866	226	-	-	-	-	-
2118507	-	0.0193	3	0	-	-	XCV1871	90/668	-	-	-
2118674	+	0.0000	9	0	XCV1872	31	-	-	XCV1871	77/835	-
2133276	-	0.0001	11	2	XCV1886	0	-	-	-	-	-
2134171	-	0.0080	5	0	-	-	XCV1887	625/802	-	-	-
2136428	-	0.0036	5	0	XCV1888	83	-	-	XCV1889	100/648	-
2136615	-	0.0356	4	0	XCV1888	270	-	-	XCV1889	87/461	-
2138655	+	0.0022	5	0	-	-	XCV1891	50/1437	-	-	-
2140105	+	0.0393	3	1	XCV1892	25	-	-	-	-	-
2149915	+	0.0034	11	0	-	-	XCV1904	174/1199	-	-	-
2153351	+	0.0000	33	2	XCV1907	111	-	-	-	-	-
2163985	+	0.0000	45	9	XCV1915	22	-	-	-	-	-
2171527	+	0.0078	4	1	-	-	-	-	XCV1920	303/563	-
2173745	-	0.0019	6	0	XCV1921	56	-	-	XCV1922	21/488	-
2173764	+	0.0177	3	1	-	-	XCV1922	40/469	XCV1921	75/1754	-
2177450	-	0.0000	7	0	XCV1924	34	-	-	-	-	-

[a], TSS position on the *Xcv* chromosome and the plasmid (pXCV183; marked with #), respectively. [b], p-value indicates the significance of the annotated TSS (see METHODS). [c], number of read starts at TSS position. [d], TSS classification and the corresponding CDS. [e], distance of the TSS to annotated start and stop codon of the corresponding CDS. [f], a TSS is classified as orphan (+) if it is neither classified as primary, internal nor antisense TSS.

TSS position[a]	Strand	p-value[b]	Library 2[c]	Library 1[c]	Primary[d] to CDS	5' UTR (bp)	Internal[d] to CDS	Distance to start/stop[e] (bp)	Antisense[d] to CDS	Distance to start/stop[e] (bp)	Orphan[f]
2177544	+	0.0048	6	1	XCV1926 XCV1925	265 20	-	-	-	-	-
2181033	+	0.0009	4	0	XCV1928	17	-	-	-	-	-
2185929	+	0.0206	3	0	-	-	-	-	XCV1931	407/186	-
2190346	-	0.0000	10	0	XCV1934	56	-	-	-	-	-
2191418	+	0.0014	4	0	-	-	-	-	XCV1936	728/54	-
2192469	+	0.0054	3	0	-	-	-	-	XCV1937	282/20	-
2201842	+	0.0206	3	0	-	-	-	-	XCV1941	2237/234	-
2221729	-	0.0018	3	0	-	-	XCV1952	1758/416	-	-	-
2226060	+	0.0031	3	0	-	-	-	-	XCV1954	1097/1173	-
2236742	-	0.0198	5	0	XCV1962	157	-	-	-	-	-
2239894	+	0.0087	3	0	-	-	XCV1963	2441/657	-	-	-
2245024	+	0.0000	19	2	-	-	XCV1969	817/13	-	-	-
2246333	+	0.0376	3	1	XCV1972	25	-	-	-	-	-
2247573	-	0.0193	3	0	-	-	-	-	XCV1973	582/893	-
2283523	-	0.0087	3	0	XCV2000	53	-	-	-	-	-
2298800	-	0.0002	6	0	XCV2016	37	-	-	-	-	-
2299637	-	0.0000	29	3	-	-	XCV2017	8/654	-	-	-
2303997	-	0.0000	15	1	XCV2022	88	-	-	-	-	-
2316851	-	0.0000	7	0	XCV2034	0	-	-	-	-	-
2317891	+	0.0138	4	1	XCV2036	22	-	-	-	-	-
2319528	+	0.0009	3	0	-	-	XCV2038	789/434	-	-	-
2331979	+	0.0000	28	2	-	-	XCV2048	60/323	-	-	-
2332427	+	0.0011	9	3	XCV2049	21	-	-	-	-	-
2332843	-	0.0400	4	0	-	-	-	-	XCV2049	395/1884	-
2332879	+	0.0100	5	0	-	-	XCV2049	431/1848	-	-	-
2335263	-	0.0000	35	5	XCV2051	40	-	-	-	-	-
2339189	-	0.0139	3	1	-	-	XCV2055	457/508	-	-	-
2339660	+	0.0000	10	1	XCV2056	270	-	-	XCV2055	14/979	-
2345698	+	0.0167	5	2	-	-	-	-	XCV2059	3686/477	-
2349615	+	0.0006	5	0	XCV2060	0	-	-	-	-	-
2350997	+	0.0135	5	2	XCV2061	171	-	-	-	-	-
2356401	+	0.0000	27	1	XCV2066	90	-	-	-	-	-
2358828	+	0.0000	30	1	XCV2069	108	XCV2068	677/108	-	-	-
2359054	+	0.0064	4	0	-	-	XCV2069	118/739	-	-	-
2359959	+	0.0145	11	6	-	-	XCV2070	92/543	-	-	-
2363433	+	0.0023	8	0	XCV2073	0	-	-	-	-	-
2363454	+	0.0023	8	0	-	-	XCV2073	21/1196	-	-	-
2368751	+	0.0309	4	0	-	-	XCV2075	2367/293	-	-	-
2373126	-	0.0028	4	0	XCV2080	22	-	-	-	-	-
2374746	+	0.0005	6	0	XCV2083	29	-	-	-	-	-
2382135	+	0.0134	3	0	XCV2090	23	-	-	XCV2089	58/1032	-
2384696	-	0.0077	4	0	-	-	-	-	XCV2092	15/1220	-
2393149	+	0.0000	29	5	XCV2097	35	-	-	-	-	-
2395930	+	0.0076	4	1	XCV2101	25	-	-	-	-	-
2396500	+	0.0018	5	0	-	-	XCV2101	545/150	-	-	-
2398280	-	0.0085	3	0	-	-	XCV2102	315/1493	-	-	-
2399163	+	0.0000	17	2	XCV2103	185	-	-	-	-	-
2400264	-	0.0193	3	0	-	-	-	-	XCV2103	916/9274	-
2411526	-	0.0190	3	0	XCV2105	0	-	-	-	-	-
2414657	-	0.0000	19	2	-	-	XCV2107	1507/1582	-	-	-
2476734	-	0.0043	4	0	-	-	-	-	XCV2155	1024/1990	-
2480236	-	0.0000	183	10	XCV2156	22	-	-	-	-	-
2480585	-	0.0122	5	1	-	-	XCV2157	769/88	-	-	-

[a], TSS position on the *Xcv* chromosome and the plasmid (pXCV183; marked with #), respectively. [b], p-value indicates the significance of the annotated TSS (see METHODS). [c], number of read starts at TSS position. [d], TSS classification and the corresponding CDS. [e], distance of the TSS to annotated start and stop codon of the corresponding CDS. [f], a TSS is classified as orphan (+) if it is neither classified as primary, internal nor antisense TSS.

TSS position[a]	Strand	p-value[b]	Library 2[c]	Library 1[c]	Primary[d] to CDS	5' UTR (bp)	Internal[d] to CDS	Distance to start/stop[e] (bp)	Antisense[d] to CDS	Distance to start/stop[e] (bp)	Orphan[f]
2480929	+	0.0000	162	23	-	-	-	-	XCV2157	425/432	-
2481679	-	0.0003	19	10	XCV2158	29	-	-	XCV2159	48/686	-
2483366	+	0.0000	12	0	-	-	XCV2161	708/938	-	-	-
2484379	-	0.0000	210	29	-	-	-	-	XCV2161 XCV2162	1721/75 5/222	-
2485206	-	0.0008	5	0	XCV2163	45	-	-	-	-	-
2485799	+	0.0003	6	0	-	-	XCV2165	73/2704	-	-	-
2490307	-	0.0101	3	0	-	-	-	-	XCV2166	1804/247	-
2495693	-	0.0096	4	0	-	-	-	-	-	-	+
2499035	-	0.0056	8	2	XCV2173	21	-	-	XCV2174	76/978	-
2511922	-	0.0000	15	0	XCV2183	98	-	-	XCV2184	18/938	-
2523234	+	0.0469	7	0	-	-	XCV2193	308/1317	-	-	-
2529317	+	0.0002	6	0	XCV2200	43	-	-	-	-	-
2531282	-	0.0031	5	1	XCV2201	124	XCV2202	2254/121	-	-	-
2532845	-	0.0013	7	2	-	-	XCV2202	691/1684	-	-	-
2533596	-	0.0000	83	7	XCV2202	60	-	-	XCV2203	83/901	-
2534689	+	0.0009	11	4	XCV2204	0	-	-	-	-	-
2536986	+	0.0000	11	2	-	-	XCV2207	27/632	-	-	-
2537657	+	0.0000	11	0	XCV2208	19	-	-	-	-	-
2540067	+	0.0068	3	0	XCV2210 XCV2211	22 228	-	-	-	-	-
2542431	+	0.0000	15	2	XCV2214	32	-	-	-	-	-
2548437	-	0.0272	6	0	XCV2215	25	-	-	-	-	-
2549014	-	0.0000	26	5	XCV2216	19	-	-	-	-	-
2552155	+	0.0018	5	0	XCV2222	125	XCV2221	185/54	-	-	-
2557002	-	0.0193	3	0	-	-	-	-	XCV2230	276/854	-
2570503	-	0.0055	3	0	-	-	XCV2237	457/1099	-	-	-
2573369	+	0.0000	65	12	XCV2240 XCV2241	14 241	-	-	XCV2239	26/895	-
2573375	-	0.0408	4	2	XCV2239	32	-	-	XCV2240	8/175	-
2580164	+	0.0023	3	0	-	-	-	-	XCV2250	1090/29	-
2580190	-	0.0020	3	0	-	-	-	-	-	-	+
2583951	+	0.0000	19	5	XCV2252	149	XCV2251	2402/144	-	-	-
2589488	+	0.0004	7	1	-	-	XCV2258	81/1415	-	-	-
2590477	+	0.0005	5	0	-	-	XCV2258	1070/426	-	-	-
2595113	-	0.0028	5	1	-	-	XCV2262	769/88	XCV2261	1384/68	-
2595457	+	0.0000	162	23	-	-	-	-	XCV2262	425/432	-
2596207	-	0.0004	19	10	XCV2263	23	XCV2264	759/146	-	-	-
2608263	+	0.0000	52	5	XCV2275	22	-	-	-	-	-
2609179	-	0.0441	4	3	-	-	-	-	XCV2275	894/170	-
2610400	-	0.0000	108	15	XCV2277	23	-	-	-	-	-
2613410	+	0.0000	33	2	XCV2280	23	-	-	-	-	-
2613918	-	0.0087	4	0	-	-	-	-	XCV2280	485/588	-
2616298	-	0.0486	4	4	-	-	XCV2283	1577/546	-	-	-
2616321	+	0.0280	7	4	-	-	-	-	XCV2283	1554/569	-
2625423	+	0.0094	3	0	-	-	-	-	XCV2290	44/796	-
2629491	+	0.0001	20	9	XCV2295	29	-	-	XCV2294	88/345	-
2630241	-	0.0000	170	21	-	-	-	-	XCV2296	425/432	-
2630585	+	0.0061	5	1	XCV2297	207	XCV2296	769/88	-	-	-
2631345	+	0.0001	8	0	-	-	-	-	-	-	+
2631738	+	0.0076	5	1	XCV2298	211	-	-	-	-	-
2634620	-	0.0125	3	0	XCV2299	37	-	-	XCV2300	68/358	-
2642688	+	0.0005	12	4	XCV2309	26	-	-	-	-	-
2645525	+	0.0293	6	0	-	-	-	-	XCV2311	235/721	-

[a], TSS position on the *Xcv* chromosome and the plasmid (pXCV183; marked with #), respectively. [b], p-value indicates the significance of the annotated TSS (see METHODS). [c], number of read starts at TSS position. [d], TSS classification and the corresponding CDS. [e], distance of the TSS to annotated start and stop codon of the corresponding CDS. [f], a TSS is classified as orphan (+) if it is neither classified as primary, internal nor antisense TSS.

TSS position[a]	Strand	p-value[b]	Library 2[c]	Library 1[c]	Primary[d] to CDS	5' UTR (bp)	Internal[d] to CDS	Distance to start/stop[e] (bp)	Antisense[d] to CDS	Distance to start/stop[e] (bp)	Orphan[f]
2646914	+	0.0085	3	0	XCV2314	1	-	-	-	-	-
2651289	-	0.0067	3	0	XCV2319	31	-	-	-	-	-
2651497	+	0.0000	151	12	-	-	XCV2320	11/780	-	-	-
2651654	+	0.0232	8	1	-	-	XCV2320	168/623	-	-	-
2652046	-	0.0009	6	1	-	-	-	-	XCV2320	560/231	-
2656549	+	0.0112	8	4	-	-	-	-	XCV2325	877/35	-
2656965	+	0.0000	12	0	-	-	-	-	XCV2325	461/381	-
2659679	+	0.0004	7	1	-	-	-	-	XCV2329	559/232	-
2660227	-	0.0000	151	14	-	-	XCV2329	11/780	XCV2330	10/501	-
2671874	+	0.0004	7	1	-	-	-	-	XCV2340	559/232	-
2672422	-	0.0000	151	14	-	-	XCV2340 XCV2341	11/780 831/50	-	-	-
2676073	+	0.0000	17	4	XCV2345	205	XCV2344	9/209	-	-	-
2680420	+	0.0140	4	1	-	-	-	-	XCV2350	539/474	-
2680631	+	0.0092	3	0	-	-	-	-	XCV2350	328/685	-
2682681	-	0.0000	1040	80	XCV2352	34	-	-	XCV2353	69/860	-
2682761	+	0.0000	151	12	-	-	XCV2353	11/780	-	-	-
2683310	-	0.0014	6	1	-	-	-	-	XCV2353	560/231	-
2684765	+	0.0021	3	0	XCV2355	119	-	-	-	-	-
2694331	+	0.0395	7	0	-	-	-	-	XCV2365	346/145	-
2740409	+	0.0004	9	1	-	-	-	-	-	-	+
2740875	+	0.0037	7	3	-	-	-	-	-	-	+
2744537	-	0.0193	3	0	XCV2419	9	-	-	XCV2420	39/401	-
2744576	+	0.0318	7	4	XCV2420	0	-	-	XCV2419	48/488	-
2745041	+	0.0026	4	0	XCV2421	0	-	-	-	-	-
2751220	-	0.0000	33	2	-	-	-	-	XCV2424	30/1013	-
2756332	+	0.0000	14	3	XCV2429	0	-	-	XCV2428	89/442	-
2764488	+	0.0011	5	0	XCV2434	22	-	-	-	-	-
2767760	-	0.0000	8	1	XCV2435	187	-	-	-	-	-
2768957	+	0.0122	8	1	-	-	-	-	-	-	+
2768989	-	0.0017	10	3	-	-	-	-	XCV2436	1026/67	-
2770178	-	0.0242	5	0	XCV2437	22	-	-	-	-	-
2779019	+	0.0351	4	0	-	-	-	-	XCV2445	298/565	-
2780755	+	0.0206	3	0	-	-	-	-	XCV2448	701/321	-
2788996	+	0.0193	3	1	XCV2461	27	-	-	-	-	-
2798003	+	0.0002	20	9	XCV2477	29	-	-	XCV2476	49/789	-
2798753	-	0.0000	170	21	-	-	-	-	XCV2478	425/432	-
2799097	+	0.0038	5	1	-	-	XCV2478	769/88	XCV2479	274/68	-
2801025	+	0.0293	5	0	-	-	XCV2482	18/1007	-	-	-
2803434	-	0.0110	4	1	-	-	XCV2484	242/957	-	-	-
2803838	-	0.0000	24	7	XCV2484	162	-	-	-	-	-
2805975	-	0.0102	3	0	-	-	XCV2486	1154/300	-	-	-
2806891	-	0.0465	3	0	-	-	XCV2486	238/1216	-	-	-
2807153	-	0.0000	12	1	XCV2486	24	-	-	-	-	-
2811009	-	0.0111	3	1	XCV2489	296	XCV2490	771/119	-	-	-
2811397	-	0.0049	3	0	-	-	XCV2490	383/507	-	-	-
2811810	-	0.0000	28	1	XCV2490	30	XCV2491	1877/27	-	-	-
2813614	+	0.0206	3	0	-	-	-	-	XCV2491 XCV2492	73/1831 1055/90	-
2817226	-	0.0044	9	9	-	-	XCV2496	1085/597	-	-	-
2818655	-	0.0000	20	5	-	-	XCV2497	522/158	-	-	-
2819280	+	0.0000	54	6	XCV2498	70	-	-	-	-	-
2819686	+	0.0000	11	0	XCV2499	31	-	-	-	-	-
2833387	-	0.0215	5	1	XCV2509	242	XCV2510	2840/135	-	-	-

[a], TSS position on the *Xcv* chromosome and the plasmid (pXCV183; marked with #), respectively. [b], p-value indicates the significance of the annotated TSS (see METHODS). [c], number of read starts at TSS position. [d], TSS classification and the corresponding CDS. [e], distance of the TSS to annotated start and stop codon of the corresponding CDS. [f], a TSS is classified as orphan (+) if it is neither classified as primary, internal nor antisense TSS.

TSS position[a]	Strand	p-value[b]	Library 2[c]	Library 1[c]	Primary[d] to CDS	5' UTR (bp)	Internal[d] to CDS	Distance to start/stop[e] (bp)	Antisense[d] to CDS	Distance to start/stop[e] (bp)	Orphan[f]
2837417	-	0.0039	7	2	XCV2511	34	-	-	-	-	-
2838360	-	0.0015	4	0	XCV2514	31	-	-	-	-	-
2838423	+	0.0000	16	1	-	-	-	-	XCV2514	94/390	-
2843234	-	0.0000	29	2	XCV2518	24	-	-	-	-	-
2843374	+	0.0051	8	0	XCV2519	0	-	-	-	-	-
2853818	+	0.0003	8	1	XCV2531	0	-	-	-	-	-
2859395	+	0.0011	5	1	-	-	XCV2535	11/1653	-	-	-
2865835	-	0.0152	5	0	XCV2539	37	-	-	-	-	-
2874392	-	0.0011	4	0	-	-	XCV2546	54/554	-	-	-
2880224	-	0.0000	36	5	XCV2551	264	-	-	-	-	-
2880816	-	0.0000	17	3	-	-	XCV2552	80/540	-	-	-
2880924	+	0.0000	9	0	XCV2553	15	-	-	XCV2552	28/648	-
2885613	-	0.0125	6	1	-	-	-	-	XCV2557 XCV2558	622/43 46/336	-
2894891	+	0.0000	45	4	XCV2568	229	-	-	XCV2567	73/690	-
2894892	+	0.0006	30	0	XCV2568	228	-	-	XCV2567	74/691	-
2897112	-	0.0028	3	0	-	-	-	-	XCV2569	681/332	-
2897115	+	0.0019	3	0	-	-	XCV2569	684/329	-	-	-
2899257	+	0.0068	3	0	-	-	XCV2571	1189/478	-	-	-
2901871	-	0.0041	5	0	-	-	XCV2574	299/417	-	-	-
2905203	-	0.0074	4	0	XCV2577	95	-	-	XCV2578	8/1033	-
2910245	+	0.0043	4	0	XCV2582	71	-	-	XCV2581	47/808	-
2915414	-	0.0000	21	2	XCV2586	0	-	-	-	-	-
2915525	+	0.0037	5	1	XCV2587	0	-	-	-	-	-
2917592	-	0.0001	6	0	XCV2588	30	-	-	-	-	-
2919213	-	0.0000	11	1	-	-	XCV2589	117/1553	XCV2590	36/1154	-
2921819	-	0.0000	349	6	XCV2592	157	-	-	-	-	-
2924141	+	0.0383	3	1	XCV2595	16	-	-	-	-	-
2927298	+	0.0210	3	0	XCV2597	4	-	-	XCV2596	94/969	-
2929160	+	0.0000	24	4	XCV2599	24	-	-	-	-	-
2931916	-	0.0004	5	0	XCV2601	29	XCV2602	1848/26	-	-	-
2933879	-	0.0041	4	0	XCV2602	115	-	-	-	-	-
2935603	-	0.0000	17	4	XCV2603	0	-	-	-	-	-
2937657	-	0.0056	4	1	XCV2605	0	-	-	XCV2606	25/1089	-
2941785	-	0.0000	30	0	XCV2608	20	-	-	-	-	-
2944333	+	0.0008	5	0	XCV2611	33	-	-	XCV2610	86/2068	-
2945265	+	0.0005	5	0	-	-	XCV2611	899/189	-	-	-
2947490	+	0.0206	3	0	-	-	-	-	XCV2612	790/1825	-
2952716	+	0.0000	22	2	-	-	XCV2618	817/13	-	-	-
2953172	-	0.0002	6	0	-	-	-	-	XCV2619	75/347	-
2953268	+	0.0000	16	2	XCV2620	279	XCV2619	21/251	-	-	-
2958333	+	0.0424	4	0	-	-	-	-	XCV2625	769/886	-
2959110	+	0.0295	3	0	XCV2626	290	-	-	XCV2625	8/1663	-
2966234	-	0.0001	8	1	XCV2630	0	-	-	-	-	-
2967021	-	0.0000	29	3	XCV2631	73	-	-	-	-	-
2970083	-	0.0000	34	8	XCV2634	71	-	-	-	-	-
2971502	-	0.0000	31	4	XCV2635	0	-	-	-	-	-
2972659	-	0.0031	4	0	-	-	XCV2636	105/962	-	-	-
2976928	+	0.0008	7	0	XCV2642	28	-	-	-	-	-
3000010	+	0.0000	45	7	XCV2661	23	-	-	-	-	-
3005175	-	0.0036	5	0	-	-	XCV2665	479/363	-	-	-
3011678	+	0.0040	3	0	XCV2671	212	XCV2670	1236/65	-	-	-
3024175	-	0.0047	6	0	-	-	XCV2674	97/5035	-	-	-
3024425	+	0.0317	3	0	XCV2675	0	-	-	-	-	-

[a], TSS position on the *Xcv* chromosome and the plasmid (pXCV183; marked with #), respectively. [b], p-value indicates the significance of the annotated TSS (see METHODS). [c], number of read starts at TSS position. [d], TSS classification and the corresponding CDS. [e], distance of the TSS to annotated start and stop codon of the corresponding CDS. [f], a TSS is classified as orphan (+) if it is neither classified as primary, internal nor antisense TSS.

TSS position[a]	Strand	p-value[b]	Library 2[c]	Library 1[c]	Primary[d] to CDS	5' UTR (bp)	Internal[d] to CDS	Distance to start/stop[e] (bp)	Antisense[d] to CDS	Distance to start/stop[e] (bp)	Orphan[f]
3027493	+	0.0024	5	1	-	-	XCV2677	1074/2093	-	-	-
3038824	-	0.0029	3	0	XCV2683	25	-	-	-	-	-
3045902	-	0.0048	5	1	XCV2688	242	XCV2689	769/88	-	-	-
3046246	+	0.0000	162	23	-	-	-	-	XCV2689	425/432	-
3046996	-	0.0002	19	10	XCV2690	29	-	-	-	-	-
3047619	-	0.0002	16	6	-	-	-	-	-	-	+
3062769	+	0.0020	3	0	XCV2703	21	-	-	-	-	-
3065332	-	0.0000	14	0	XCV2704	0	XCV2705	110/0	-	-	-
3068986	-	0.0043	6	2	XCV2709	24	XCV2710	596/24	-	-	-
3069778	+	0.0000	10	1	XCV2711	43	-	-	-	-	-
3075487	+	0.0184	4	0	-	-	XCV2715	2113/2521	-	-	-
3078655	+	0.0001	7	0	XCV2716	22	-	-	-	-	-
3087900	+	0.0000	24	5	-	-	XCV2723	817/13	-	-	-
3093405	+	0.0044	4	1	-	-	-	-	XCV2725	892/1999	-
3094186	-	0.0004	6	0	-	-	XCV2725	111/2780	-	-	-
3094539	-	0.0000	8	0	XCV2725	242	-	-	-	-	-
3095807	-	0.0025	3	0	-	-	XCV2726	769/1048	-	-	-
3103362	-	0.0033	3	0	-	-	XCV2729	22/4333	-	-	-
3106814	-	0.0000	16	1	XCV2730	214	XCV2731	164/84	-	-	-
3108818	-	0.0114	3	0	-	-	XCV2732	620/1734	-	-	-
3112840	+	0.0171	4	1	-	-	-	-	XCV2736	959/672	-
3113638	+	0.0206	3	0	-	-	-	-	XCV2736	161/1470	-
3129258	-	0.0000	71	9	XCV2748	0	-	-	-	-	-
3130321	+	0.0015	5	0	-	-	XCV2750	16/1636	-	-	-
3142239	+	0.0017	3	0	XCV2763	4	-	-	-	-	-
3143314	+	0.0000	47	7	XCV2764	19	-	-	-	-	-
3151089	-	0.0019	4	0	-	-	XCV2770	46/1018	-	-	-
3151886	+	0.0278	4	1	-	-	-	-	XCV2771 XCV2772	11/616 838/8	-
3151905	-	0.0000	17	2	XCV2771	30	XCV2772	819/11	-	-	-
3166096	-	0.0024	3	0	-	-	XCV2786	1363/58	-	-	-
3168287	-	0.0000	20	1	XCV2787	133	-	-	-	-	-
3174731	-	0.0001	7	1	XCV2794	263	XCV2795	1753/148	-	-	-
3184488	-	0.0001	10	2	-	-	-	-	XCV2800	857/1221	-
3193300	+	0.0000	20	2	-	-	XCV2806	817/13	-	-	-
3196040	-	0.0001	12	1	XCV2809	257	-	-	-	-	-
3198093	-	0.0001	8	0	XCV2811	25	-	-	-	-	-
3201040	-	0.0005	7	1	XCV2812	90	-	-	XCV2813	21/518	-
3201949	+	0.0156	4	1	-	-	-	-	XCV2814	203/201	-
3202518	+	0.0117	3	0	-	-	-	-	XCV2815	3413/330	-
3205497	-	0.0050	3	0	-	-	XCV2815	434/3309	-	-	-
3206053	+	0.0226	4	1	-	-	-	-	XCV2816	838/8	-
3206074	-	0.0000	15	2	XCV2815	143	XCV2816	817/13	-	-	-
3207206	-	0.0292	6	2	XCV2817	16	-	-	-	-	-
3209078	+	0.0000	12	0	-	-	-	-	XCV2820	266/210	-
3209434	-	0.0000	28	2	XCV2820	90	XCV2821	425/90	-	-	-
3209880	-	0.0000	14	0	XCV2821	21	-	-	-	-	-
3211387	-	0.0000	69	15	XCV2822	18	-	-	-	-	-
3214823	-	0.0020	3	0	XCV2823	123	-	-	-	-	-
3217042	-	0.0146	3	1	XCV2826	57	-	-	XCV2827	64/639	-
3219491	-	0.0000	15	1	XCV2829	21	-	-	XCV2830	38/304	-
3219506	+	0.0001	7	0	XCV2831 XCV2830	289 23	-	-	XCV2829	36/536	-
3223002	-	0.0142	3	0	-	-	XCV2833	1321/790	-	-	-

[a], TSS position on the *Xcv* chromosome and the plasmid (pXCV183; marked with #), respectively. [b], p-value indicates the significance of the annotated TSS (see METHODS). [c], number of read starts at TSS position. [d], TSS classification and the corresponding CDS. [e], distance of the TSS to annotated start and stop codon of the corresponding CDS. [f], a TSS is classified as orphan (+) if it is neither classified as primary, internal nor antisense TSS.

TSS position[a]	Strand	p-value[b]	Library 2[c]	Library 1[c]	Primary[d] to CDS	5' UTR (bp)	Internal[d] to CDS	Distance to start/stop[e] (bp)	Antisense[d] to CDS	Distance to start/stop[e] (bp)	Orphan[f]
3223483	-	0.0042	4	0	-	-	XCV2833	840/1271	-	-	-
3227994	-	0.0041	6	1	-	-	XCV2837	1172/1539	-	-	-
3229766	-	0.0087	3	0	-	-	XCV2838	1007/501	-	-	-
3231114	-	0.0002	6	1	-	-	XCV2839	254/333	-	-	-
3233114	-	0.0008	3	0	-	-	XCV2840	192/1268	-	-	-
3239149	-	0.0102	4	1	-	-	XCV2846	433/655	-	-	-
3247561	-	0.0000	10	1	XCV2854	185	XCV2855	604/148	-	-	-
3248192	-	0.0005	9	0	XCV2855	27	-	-	XCV2856	86/883	-
3258473	+	0.0169	3	1	-	-	-	-	XCV2864	21/905	-
3259835	-	0.0036	5	1	XCV2865	289	XCV2866	769/88	-	-	-
3260179	+	0.0000	162	23	-	-	-	-	XCV2866	425/432	-
3260929	-	0.0004	19	10	XCV2867	29	XCV2868	266/75	-	-	-
3262486	-	0.0007	10	3	XCV2869	31	-	-	XCV2870	74/997	-
3267661	-	0.0000	16	3	-	-	XCV2874	105/1976	-	-	-
3270081	-	0.0406	4	2	-	-	XCV2876	462/647	-	-	-
3272560	-	0.0050	4	1	XCV2878	23	-	-	XCV2879	58/705	-
3272598	+	0.0103	4	1	XCV2879	20	-	-	XCV2878	61/984	-
3275968	-	0.0158	6	3	XCV2881	21	-	-	-	-	-
3276220	+	0.0041	5	1	XCV2882	20	-	-	-	-	-
3278748	+	0.0019	5	1	XCV2884	23	-	-	-	-	-
3284009	-	0.0058	3	0	XCV2887	42	-	-	-	-	-
3285246	+	0.0000	12	1	XCV2890	23	-	-	-	-	-
3285826	+	0.0001	8	1	XCV2891	21	-	-	-	-	-
3290997	-	0.0001	12	2	-	-	-	-	-	-	+
3291196	+	0.0042	4	1	XCV2895	107	-	-	-	-	-
3292629	+	0.0216	3	0	-	-	XCV2896	652/1678	-	-	-
3296402	+	0.0003	5	0	XCV2898	28	-	-	-	-	-
3298935	+	0.0001	7	0	-	-	XCV2899	236/2115	-	-	-
3301211	+	0.0002	8	1	XCV2900	27	-	-	-	-	-
3304175	-	0.0012	14	0	XCV2901	18	-	-	-	-	-
3304605	-	0.0124	3	0	XCV2902	0	-	-	-	-	-
3309674	+	0.0010	4	0	XCV2907	27	-	-	XCV2906	83/559	-
3309679	-	0.0000	34	2	XCV2906	88	-	-	XCV2907	22/561	-
3310327	+	0.0000	35	2	XCV2908	26	-	-	-	-	-
3311814	-	0.0012	7	2	XCV2909	23	-	-	-	-	-
3316609	+	0.0010	4	0	-	-	XCV2914	211/1138	-	-	-
3317782	+	0.0000	48	8	XCV2916 XCV2915	262 0	-	-	-	-	-
3320950	-	0.0001	12	3	XCV2919	23	-	-	-	-	-
3323956	-	0.0145	3	1	-	-	XCV2923	1168/274	-	-	-
3325124	-	0.0022	8	3	XCV2923	0	-	-	XCV2924	51/848	-
3332244	+	0.0000	19	1	XCV2929	17	-	-	-	-	-
3334516	-	0.0000	17	1	XCV2931	0	-	-	-	-	-
3336764	-	0.0000	16	0	XCV2933	20	-	-	-	-	-
3339362	-	0.0000	14	0	XCV2936	0	-	-	-	-	-
3341585	-	0.0005	9	2	-	-	XCV2938	1132/1663	-	-	-
3343425	+	0.0027	7	0	XCV2940	28	-	-	-	-	-
3348398	+	0.0001	12	3	XCV2944	173	-	-	-	-	-
3349667	-	0.0231	3	0	-	-	-	-	XCV2944	1096/1102	-
3355947	+	0.0211	3	0	-	-	XCV2950	704/711	-	-	-
3358746	+	0.0002	11	3	XCV2955	22	XCV2954	148/1	-	-	-
3373502	-	0.0004	6	0	XCV2966	25	-	-	-	-	-
3377499	+	0.0000	8	0	XCV2970	0	-	-	-	-	-
3379701	-	0.0001	7	1	-	-	XCV2971	37/1087	-	-	-

[a], TSS position on the *Xcv* chromosome and the plasmid (pXCV183; marked with #), respectively. [b], p-value indicates the significance of the annotated TSS (see METHODS). [c], number of read starts at TSS position. [d], TSS classification and the corresponding CDS. [e], distance of the TSS to annotated start and stop codon of the corresponding CDS. [f], a TSS is classified as orphan (+) if it is neither classified as primary, internal nor antisense TSS.

TSS position[a]	Strand	p-value[b]	Library 2[c]	Library 1[c]	Primary[d] to CDS	5' UTR (bp)	Internal[d] to CDS	Distance to start/stop[e] (bp)	Antisense[d] to CDS	Distance to start/stop[e] (bp)	Orphan[f]
3380658	-	0.0147	4	0	-	-	XCV2972	39/719	-	-	-
3381360	-	0.0035	3	0	-	-	XCV2973	1843/559	-	-	-
3385409	-	0.0000	12	1	-	-	XCV2974	46/619	XCV2975	46/619	-
3388971	-	0.0035	5	0	XCV2981	23	-	-	-	-	-
3390399	+	0.0035	3	0	XCV2983	0	-	-	-	-	-
3395819	-	0.0027	5	2	XCV2987	82	-	-	-	-	-
3402000	+	0.0000	13	3	-	-	XCV2992	179/66	-	-	-
3403164	-	0.0051	3	0	-	-	-	-	XCV2993	743/1092	-
3408906	-	0.0127	4	0	-	-	-	-	-	-	+
3410761	+	0.0000	15	3	XCV3000	6	-	-	-	-	-
3413795	-	0.0049	5	0	-	-	XCV3002	256/1246	-	-	-
3418545	+	0.0055	4	0	XCV3005	16	-	-	-	-	-
3424210	-	0.0016	5	0	-	-	XCV3010	63/599	-	-	-
3426246	+	0.0002	10	2	XCV3013	178	-	-	-	-	-
3429740	+	0.0009	4	0	XCV3016	0	-	-	-	-	-
3449971	-	0.0042	4	1	XCV3029	27	-	-	XCV3030	76/786	-
3450047	+	0.0067	4	1	XCV3030	0	-	-	-	-	-
3452981	+	0.0000	14	1	XCV3033	110	-	-	-	-	-
3455491	+	0.0025	6	0	XCV3036	21	-	-	-	-	-
3459115	-	0.0105	10	5	-	-	XCV3037	8/693	-	-	-
3473908	-	0.0000	57	0	XCV3049	55	-	-	-	-	-
3474110	+	0.0001	9	1	XCV3050	19	-	-	-	-	-
3477013	-	0.0179	3	1	-	-	XCV3053	848/555	-	-	-
3478034	-	0.0001	6	0	XCV3053	173	XCV3054	1163/189	-	-	-
3482294	-	0.0016	4	0	XCV3056	49	-	-	-	-	-
3485108	-	0.0003	5	0	-	-	XCV3059	65/621	-	-	-
3485693	-	0.0001	11	0	XCV3060	23	-	-	-	-	-
3487971	-	0.0010	7	0	XCV3063	0	-	-	-	-	-
3488858	+	0.0200	8	0	-	-	XCV3065	12/893	-	-	-
3491956	-	0.0038	3	0	-	-	XCV3067	280/847	-	-	-
3493428	-	0.0000	11	1	XCV3068	42	-	-	-	-	-
3496031	+	0.0002	13	8	-	-	XCV3071	421/346	-	-	-
3496288	+	0.0000	13	3	XCV3072	199	XCV3071	678/89	-	-	-
3497562	+	0.0024	3	0	XCV3074	22	-	-	-	-	-
3498161	+	0.0052	7	2	XCV3075	21	-	-	-	-	-
3505699	+	0.0000	18	4	XCV3083	0	-	-	-	-	-
3518387	-	0.0000	994	98	XCV3096	0	-	-	-	-	-
3525219	-	0.0000	10	0	XCV3103	14	-	-	-	-	-
3525280	+	0.0031	3	0	XCV3104	57	-	-	XCV3103	75/1184	-
3528942	+	0.0037	3	0	-	-	XCV3108	507/173	-	-	-
3531766	-	0.0000	12	1	XCV3111	281	XCV3112	8/246	XCV3113	32/1030	-
3533944	+	0.0020	4	1	XCV3116	2	XCV3115	551/21	-	-	-
3534466	+	0.0045	3	1	XCV3117	40	XCV3116	520/37	-	-	-
3536664	+	0.0000	95	12	XCV3119	129	XCV3118	1356/38	-	-	-
3539930	+	0.0284	24	15	XCV3123	0	-	-	-	-	-
3539954	+	0.0062	11	0	-	-	XCV3123	24/365	-	-	-
3544261	+	0.0001	10	1	XCV3127	23	-	-	-	-	-
3545188	-	0.0029	5	1	-	-	XCV3128	970/43	-	-	-
3551148	-	0.0015	4	0	XCV3131	23	-	-	-	-	-
3557788	-	0.0203	4	0	-	-	XCV3138	1006/385	-	-	-
3573400	-	0.0028	8	3	XCV3149	21	-	-	XCV3150	79/1449	-
3574632	-	0.0019	4	0	-	-	-	-	XCV3150	1153/217	-
3575831	-	0.0011	3	0	-	-	XCV3152	752/411	-	-	-
3581398	-	0.0058	4	1	XCV3159	23	-	-	-	-	-

[a], TSS position on the *Xcv* chromosome and the plasmid (pXCV183; marked with #), respectively. [b], p-value indicates the significance of the annotated TSS (see METHODS). [c], number of read starts at TSS position. [d], TSS classification and the corresponding CDS. [e], distance of the TSS to annotated start and stop codon of the corresponding CDS. [f], a TSS is classified as orphan (+) if it is neither classified as primary, internal nor antisense TSS.

TSS position[a]	Strand	p-value[b]	Library 2[c]	Library 1[c]	Primary[d] to CDS	5' UTR (bp)	Internal[d] to CDS	Distance to start/stop[e] (bp)	Antisense[d] to CDS	Distance to start/stop[e] (bp)	Orphan[f]
3583485	+	0.0094	5	1	XCV3163 XCV3162	2690	-	-	-	-	-
3590242	-	0.0437	7	0	XCV3165	21	-	-	-	-	-
3601774	-	0.0055	4	1	XCV3168	22	-	-	XCV3169	75/734	-
3601849	+	0.0033	8	3	XCV3169	0	-	-	XCV3168	97/3669	-
3603357	+	0.0017	3	0	-	-	XCV3171	276/290	-	-	-
3610521	-	0.0000	53	4	XCV3177	188	-	-	-	-	-
3616498	-	0.0000	26	5	XCV3182	24	XCV3183	982/61	-	-	-
3620330	+	0.0000	31	3	XCV3187	89	-	-	-	-	-
3620973	+	0.0033	9	13	-	-	XCV3187	554/2343	-	-	-
3622644	-	0.0072	3	0	-	-	-	-	XCV3187	2225/672	-
3622778	+	0.0000	6	3	-	-	XCV3187	2359/538	-	-	-
3624042	-	0.0000	14	2	-	-	XCV3188	42/605	-	-	-
3628252	-	0.0020	4	0	XCV3191	23	-	-	-	-	-
3629839	-	0.0000	9	0	XCV3192	69	-	-	-	-	-
3630079	+	0.0000	41	5	XCV3193	25	-	-	-	-	-
3630391	+	0.0204	4	0	-	-	XCV3193	287/1089	-	-	-
3634081	-	0.0009	6	1	XCV3195	36	-	-	-	-	-
3634438	-	0.0010	7	0	-	-	XCV3196	865/241	-	-	-
3635324	-	0.0001	10	0	XCV3196	21	-	-	-	-	-
3639271	+	0.0000	16	2	XCV3202	0	-	-	XCV3201	33/587	-
3639290	-	0.0003	10	2	XCV3201	52	-	-	XCV3202	19/1396	-
3643121	+	0.0000	43	3	-	-	-	-	XCV3205	79/246	-
3672452	-	0.0000	9	0	-	-	XCV3222	757/304	-	-	-
3673257	-	0.0074	4	1	XCV3222	48	-	-	-	-	-
3680238	-	0.0221	4	1	-	-	XCV3228	238/967	-	-	-
3682598	-	0.0019	6	1	-	-	XCV3229	5082/2132	-	-	-
3683097	-	0.0018	3	0	-	-	XCV3229	4583/2631	-	-	-
3684784	-	0.0103	3	1	-	-	XCV3229	2896/4318	-	-	-
3691234	-	0.0115	16	7	XCV3233	53	-	-	-	-	-
3691397	+	0.0047	4	0	XCV3234	21	-	-	-	-	-
3698089	-	0.0219	3	1	-	-	XCV3240	342/179	-	-	-
3700913	-	0.0008	7	1	XCV3241	21	-	-	-	-	-
3710992	+	0.0001	11	0	XCV3249	14	-	-	-	-	-
3713718	+	0.0000	10	1	XCV3251	45	-	-	-	-	-
3721917	-	0.0320	3	1	XCV3253	44	-	-	XCV3254	59/604	-
3725397	+	0.0001	7	0	-	-	XCV3257	7/421	-	-	-
3727405	+	0.0000	21	1	-	-	-	-	-	-	+
3728571	+	0.0000	32	1	XCV3260	18	-	-	-	-	-
3732545	+	0.0081	3	0	XCV3264	21	-	-	XCV3263	81/704	-
3735352	+	0.0000	13	1	XCV3268	26	-	-	XCV3267	23/1255	-
3738534	-	0.0000	177	26	XCV3271	42	-	-	-	-	-
3739927	-	0.0263	4	1	XCV3272	58	-	-	-	-	-
3744037	-	0.0006	6	1	XCV3277	0	-	-	-	-	-
3746306	-	0.0008	5	0	-	-	XCV3279	796/1105	-	-	-
3748830	-	0.0020	3	0	XCV3281	265	XCV3282	594/131	-	-	-
3749453	-	0.0000	22	4	XCV3282	29	-	-	-	-	-
3750836	-	0.0003	6	0	XCV3284	22	-	-	XCV3285	23/435	-
3763733	-	0.0193	3	0	XCV3293	20	-	-	XCV3294	81/953	-
3765356	-	0.0062	4	0	XCV3295	19	-	-	-	-	-
3767403	+	0.0000	14	2	XCV3297	91	-	-	XCV3296	41/2002	-
3793853	-	0.0018	4	0	XCV3315	51	XCV3316	466/55	-	-	-
3804274	-	0.0017	3	0	-	-	XCV3324	604/1978	-	-	-
3819837	+	0.0000	21	0	XCV3337	152	-	-	-	-	-

[a], TSS position on the *Xcv* chromosome and the plasmid (pXCV183; marked with #), respectively. [b], p-value indicates the significance of the annotated TSS (see METHODS). [c], number of read starts at TSS position. [d], TSS classification and the corresponding CDS. [e], distance of the TSS to annotated start and stop codon of the corresponding CDS. [f], a TSS is classified as orphan (+) if it is neither classified as primary, internal nor antisense TSS.

TSS position[a]	Strand	p-value[b]	Library 2[c]	Library 1[c]	Primary[d] to CDS	5' UTR (bp)	Internal[d] to CDS	Distance to start/stop[e] (bp)	Antisense[d] to CDS	Distance to start/stop[e] (bp)	Orphan[f]
3827398	-	0.0129	3	1	-	-	XCV3342	613/379	-	-	-
3828101	+	0.0000	14	0	XCV3343	32	-	-	XCV3342	90/1082	-
3832104	-	0.0006	5	0	XCV3345	170	XCV3346	132/155	-	-	-
3834377	+	0.0206	3	0	-	-	-	-	XCV3349	459/707	-
3837008	+	0.0000	18	0	-	-	XCV3351	60/1391	-	-	-
3840372	-	0.0000	59	3	XCV3352	22	-	-	-	-	-
3840688	+	0.0000	27	0	-	-	-	-	XCV3353	224/201	-
3841040	-	0.0000	41	9	XCV3353	128	-	-	-	-	-
3841288	-	0.0375	6	0	-	-	-	-	XCV3354	34/1225	-
3841563	+	0.0014	3	0	-	-	XCV3354	309/950	-	-	-
3842377	+	0.0037	4	1	XCV3355	146	XCV3354	1123/136	-	-	-
3845146	-	0.0280	3	1		-	-	-	XCV3357	953/84	-
3850198	+	0.0000	23	2	-	-	XCV3360	817/13	-	-	-
3854811	-	0.0000	12	1	XCV3366	29	-	-	-	-	-
3859554	-	0.0002	6	0	XCV3370	25	-	-	-	-	-
3861697	-	0.0004	6	0	-	-	-	-	-	-	+
3862413	+	0.0057	7	3	XCV3376	0	-	-	-	-	-
3866613	-	0.0029	3	0	XCV3380	25	-	-	-	-	-
3868645	-	0.0063	3	0	-	-	XCV3381	1945/2014	-	-	-
3871397	-	0.0028	5	1	XCV3382	115	XCV3383	769/88	-	-	-
3871741	+	0.0000	162	23	-	-	-	-	XCV3383	425/432	-
3872491	-	0.0072	19	10	XCV3384	29	-	-	XCV3385	17/252	-
3872641	-	0.0000	20	0	XCV3384	179	-	-	XCV3385	167/102	-
3873924	+	0.0000	20	2	-	-	XCV3387	817/13	XCV3388	285/46	-
3877920	-	0.0000	26	2	XCV3391	247	-	-	-	-	-
3880139	-	0.0000	52	3	-	-	XCV3392	46/1567	XCV3393	35/703	-
3880635	-	0.0105	3	0	-	-	-	-	XCV3393	461/207	-
3884056	+	0.0001	20	9	XCV3397	29	XCV3396	1586/75	-	-	-
3884806	-	0.0000	170	21	-	-	-	-	XCV3398	425/432	-
3885150	+	0.0048	5	1	-	-	XCV3398	769/88	-	-	-
3886210	+	0.0000	17	6	XCV3399	173	-	-	-	-	-
3886493	-	0.0037	4	0	-	-	-	-	XCV3399	110/1965	-
3897296	-	0.0000	42	7	XCV3408	24	-	-	-	-	-
3897851	+	0.0000	162	23	-	-	-	-	XCV3409	425/432	-
3898601	-	0.0002	19	10	XCV3410	29	-	-	-	-	-
3899941	-	0.0104	5	0	-	-	XCV3413	1081/61	XCV3412	888/64	-
3903440	+	0.0000	13	1	XCV3416 XCV3417	0 250	-	-	-	-	-
3908082	+	0.0002	7	0	XCV3421	27	-	-	-	-	-
3909587	+	0.0055	3	0	XCV3424	36	-	-	-	-	-
3910153	+	0.0074	4	1	XCV3425	25	-	-	-	-	-
3920060	-	0.0008	8	1	XCV3430	142	-	-	XCV3431	77/2038	-
3927580	+	0.0000	15	1	-	-	-	-	-	-	+
3931813	-	0.0000	14	1	XCV3437	22	-	-	-	-	-
3935639	+	0.0044	4	0	-	-	XCV3440	1960/1993	-	-	-
3939554	-	0.0018	5	1	-	-	-	-	XCV3442	1004/306	-
3944243	-	0.0036	4	0	XCV3444	21	-	-	-	-	-
3944604	-	0.0033	3	0	-	-	-	-	-	-	+
3956443	-	0.0040	8	3	-	-	XCV3452	111/434	-	-	-
3956726	+	0.0022	5	0	-	-	XCV3453	144/1214	-	-	-
3962678	+	0.0000	9	0	XCV3460	19	-	-	-	-	-
3963400	+	0.0000	26	4	XCV3461	0	-	-	-	-	-
3964399	-	0.0202	3	0	-	-	XCV3462	760/241	-	-	-
3965186	-	0.0002	5	0	XCV3462	27	-	-	-	-	-

[a], TSS position on the *Xcv* chromosome and the plasmid (pXCV183; marked with #), respectively. [b], p-value indicates the significance of the annotated TSS (see METHODS). [c], number of read starts at TSS position. [d], TSS classification and the corresponding CDS. [e], distance of the TSS to annotated start and stop codon of the corresponding CDS. [f], a TSS is classified as orphan (+) if it is neither classified as primary, internal nor antisense TSS.

TSS position[a]	Strand	p-value[b]	Library 2[c]	Library 1[c]	Primary[d] to CDS	5' UTR (bp)	Internal[d] to CDS	Distance to start/stop[e] (bp)	Antisense[d] to CDS	Distance to start/stop[e] (bp)	Orphan[f]
3969443	-	0.0020	3	0	XCV3465	258	-	-	-	-	-
3974418	-	0.0003	6	0	XCV3468	65	-	-	-	-	-
3975609	-	0.0000	36	5	XCV3469	45	-	-	-	-	-
3976811	+	0.0112	5	2	XCV3471	26	-	-	-	-	-
3977787	-	0.0001	9	1	-	-	-	-	XCV3472	209/402	-
3978409	+	0.0002	8	0	-	-	XCV3473	141/812	-	-	-
3980115	+	0.0007	5	0	XCV3475	23	-	-	-	-	-
3984582	-	0.0403	3	1	-	-	XCV3481	71/381	-	-	-
3991134	+	0.0046	3	0	XCV3486	0	-	-	-	-	-
3999749	-	0.0225	3	0	-	-	XCV3491	21/1304	-	-	-
4001676	+	0.0155	3	1	-	-	-	-	XCV3492	620/1125	-
4007577	-	0.0000	28	2	XCV3496	0	-	-	-	-	-
4008425	-	0.0006	6	3	-	-	XCV3497	1539/407	-	-	-
4012132	-	0.0017	6	0	XCV3500	174	XCV3501	884/171	-	-	-
4013200	+	0.0010	6	1	XCV3502	29	-	-	-	-	-
4018650	-	0.0000	242	26	XCV3505 XCV3506	272 27	-	-	-	-	-
4018651	-	0.0025	76	7	XCV3505 XCV3506	273 28	-	-	-	-	-
4019632	+	0.0011	8	2	-	-	-	-	XCV3507	1/934	-
4020256	-	0.0157	3	0	-	-	XCV3508	1653/497	-	-	-
4022331	-	0.0000	11	1	XCV3509	21	-	-	-	-	-
4024561	-	0.0001	9	1	XCV3510	0	-	-	-	-	-
4025035	-	0.0000	25	2	XCV3511	47	-	-	-	-	-
4026894	-	0.0018	7	2	XCV3513	22	-	-	XCV3514	88/810	-
4027759	-	0.0224	3	0	-	-	-	-	XCV3514 XCV3515	777/55 32/334	-
4037253	+	0.0022	14	7	XCV3525	218	XCV3524	207/218	-	-	-
4037865	+	0.0000	2627	442	-	-	-	-	-	-	+
4037913	-	0.0000	36	17	-	-	-	-	-	-	+
4041002	+	0.0024	3	0	-	-	XCV3530	147/842	-	-	-
4042199	+	0.0120	3	0	-	-	XCV3532	28/691	XCV3531	5/204	-
4046356	+	0.0469	7	0	-	-	-	-	XCV3534	558/2621	-
4061390	-	0.0014	4	0	XCV3544	28	-	-	XCV3545	45/914	-
4064047	-	0.0329	5	0	-	-	XCV3547	69/656	-	-	-
4064199	+	0.0071	5	1	XCV3548	0	-	-	XCV3547	83/808	-
4064434	+	0.0046	4	0	-	-	XCV3548	235/1849	-	-	-
4068348	+	0.0000	18	2	-	-	-	-	XCV3551	110/483	-
4069816	-	0.0013	7	1	XCV3552	0	-	-	-	-	-
4069950	+	0.0000	35	10	-	-	-	-	-	-	+
4070819	-	0.0003	7	0	XCV3553	78	-	-	XCV3554	82/1335	-
4072489	+	0.0000	27	2	XCV3556 XCV3555	240 94	-	-	-	-	-
4084692	-	0.0083	3	0	-	-	-	-	XCV3566	176/1173	-
4086583	-	0.0069	6	2	XCV3567	29	-	-	-	-	-
4089769	+	0.0000	44	3	XCV3570	32	-	-	-	-	-
4090439	+	0.0085	4	1	XCV3571	45	-	-	-	-	-
4098408	+	0.0006	5	0	-	-	-	-	XCV3574	31/2493	-
4100738	-	0.0000	25	1	XCV3575	184	-	-	-	-	-
4107864	-	0.0154	3	0	XCV3578	29	-	-	-	-	-
4108094	+	0.0000	30	8	-	-	-	-	-	-	+
4109000	+	0.0025	4	0	-	-	XCV3579	513/485	-	-	-
4112499	+	0.0105	3	0	XCV3583	186	XCV3582	929/186	-	-	-
4117962	-	0.0038	4	0	XCV3586	43	-	-	XCV3587	78/959	-

[a], TSS position on the *Xcv* chromosome and the plasmid (pXCV183; marked with #), respectively. [b], p-value indicates the significance of the annotated TSS (see METHODS). [c], number of read starts at TSS position. [d], TSS classification and the corresponding CDS. [e], distance of the TSS to annotated start and stop codon of the corresponding CDS. [f], a TSS is classified as orphan (+) if it is neither classified as primary, internal nor antisense TSS.

TSS position[a]	Strand	p-value[b]	Library 2[c]	Library 1[c]	Primary[d] to CDS	5' UTR (bp)	Internal[d] to CDS	Distance to start/stop[e] (bp)	Antisense[d] to CDS	Distance to start/stop[e] (bp)	Orphan[f]
4122387	+	0.0009	8	2	XCV3589	0	-	-	-	-	-
4123145	+	0.0016	5	0	XCV3590	24	-	-	-	-	-
4130785	-	0.0035	9	3	XCV3596 XCV3595	28 236	-	-	-	-	-
4137131	-	0.0000	10	0	XCV3598	28	-	-	-	-	-
4160234	+	0.0144	3	0	-	-	XCV3617	377/2082	-	-	-
4165173	-	0.0193	3	0	XCV3619	34	-	-	-	-	-
4167914	+	0.0000	10	1	XCV3622	24	-	-	-	-	-
4171902	-	0.0026	5	1	-	-	XCV3625	15/773	XCV3626	61/984	-
4171936	+	0.0001	12	3	XCV3626	27	-	-	XCV3625	19/807	-
4176511	+	0.0483	3	0	-	-	XCV3630	293/984	-	-	-
4183458	-	0.0153	5	0	XCV3636	21	-	-	-	-	-
4199423	-	0.0272	6	0	XCV3647	2	-	-	-	-	-
4205166	-	0.0363	4	0	XCV3651	104	-	-	-	-	-
4213401	-	0.0000	14	3	-	-	XCV3658	2181/95	-	-	-
4214082	-	0.0152	3	1	-	-	XCV3658	1500/776	-	-	-
4223074	-	0.0022	3	0	-	-	XCV3668	1292/432	-	-	-
4224385	-	0.0043	3	0	XCV3668	19	-	-	-	-	-
4245230	-	0.0001	8	1	XCV3675	24	-	-	-	-	-
4247726	-	0.0193	3	0	-	-	XCV3678	165/131	-	-	-
4258733	+	0.0037	4	0	XCV3686	283	-	-	XCV3685	43/2833	-
4259551	+	0.0000	45	4	-	-	XCV3686	535/1456	-	-	-
4264986	-	0.0000	16	5	XCV3690	0	-	-	-	-	-
4265843	+	0.0000	23	1	XCV3692	21	-	-	-	-	-
4267656	+	0.0000	13	1	XCV3693	46	-	-	-	-	-
4269632	+	0.0001	7	0	-	-	XCV3695	102/1124	-	-	-
4270597	+	0.0017	4	0	XCV3696	248	XCV3695	1067/159	-	-	-
4274499	+	0.0003	8	1	XCV3700	27	-	-	XCV3699	89/685	-
4277595	-	0.0000	23	6	XCV3702	26	-	-	-	-	-
4279066	+	0.0042	5	2	XCV3704	31	-	-	-	-	-
4282216	-	0.0002	7	0	XCV3705	260	-	-	-	-	-
4283564	-	0.0041	3	1	-	-	XCV3707	189/365	-	-	-
4283579	+	0.0018	5	0	-	-	-	-	XCV3707	174/380	-
4284080	-	0.0011	5	0	-	-	XCV3708	557/327	-	-	-
4284757	-	0.0000	13	2	XCV3708	120	XCV3709	991/61	-	-	-
4285797	-	0.0017	5	0	XCV3709	49	-	-	-	-	-
4287776	+	0.0000	20	2	-	-	-	-	-	-	+
4292574	-	0.0066	3	0	-	-	-	-	XCV3716	871/1378	-
4293883	+	0.0001	7	0	-	-	XCV3716	2180/69	-	-	-
4296181	-	0.0047	6	2	XCV3718	82	-	-	-	-	-
4309263	-	0.0000	12	1	XCV3726	32	-	-	-	-	-
4309871	-	0.0055	9	3	XCV3727	186	-	-	-	-	-
4311306	+	0.0001	9	2	XCV3730	100	-	-	XCV3729	1/379	-
4311336	-	0.0000	8	0	XCV3729	29	-	-	XCV3730	70/1200	-
4314643	-	0.0000	12	0	XCV3732	23	-	-	-	-	-
4314757	+	0.0008	7	2	XCV3733	241	-	-	-	-	-
4316512	-	0.0299	3	0	-	-	-	-	XCV3734	10/360	-
4317032	+	0.0000	12	2	-	-	XCV3735	82/367	-	-	-
4318226	-	0.0018	5	1	XCV3736	37	-	-	-	-	-
4318420	+	0.0010	7	1	XCV3737	0	-	-	-	-	-
4321777	+	0.0001	8	1	XCV3741	33	-	-	-	-	-
4323846	-	0.0166	3	1	-	-	XCV3743	672/32	-	-	-
4326863	-	0.0448	3	1	XCV3744	182	-	-	-	-	-
4328091	+	0.0001	13	3	-	-	XCV3746	16/439	-	-	-

[a], TSS position on the *Xcv* chromosome and the plasmid (pXCV183; marked with #), respectively. [b], p-value indicates the significance of the annotated TSS (see METHODS). [c], number of read starts at TSS position. [d], TSS classification and the corresponding CDS. [e], distance of the TSS to annotated start and stop codon of the corresponding CDS. [f], a TSS is classified as orphan (+) if it is neither classified as primary, internal nor antisense TSS.

TSS position[a]	Strand	p-value[b]	Library 2[c]	Library 1[c]	Primary[d] to CDS	5' UTR (bp)	Internal[d] to CDS	Distance to start/stop[e] (bp)	Antisense[d] to CDS	Distance to start/stop[e] (bp)	Orphan[f]
4339048	+	0.0200	4	0	-	-	-	-	XCV3754	365/1461	-
4339435	-	0.0006	9	3	XCV3754	22	-	-	-	-	-
4349332	-	0.0151	3	0	XCV3762	23	-	-	XCV3763	92/502	-
4350995	-	0.0003	6	0	XCV3765	72	-	-	-	-	-
4358104	-	0.0422	3	1	-	-	XCV3773	121/676	-	-	-
4358669	-	0.0000	13	2	XCV3774	86	-	-	-	-	-
4364685	-	0.0000	45	0	XCV3779	31	-	-	-	-	-
4366433	+	0.0296	3	0	XCV3782	40	-	-	-	-	-
4367233	-	0.0359	11	0	-	-	-	-	XCV3782 XCV3783	760/77 89/694	-
4367263	+	0.0001	6	0	XCV3783	59	-	-	-	-	-
4367880	-	0.0082	8	1	-	-	-	-	XCV3783	558/47	-
4369125	-	0.0193	3	0	XCV3784	23	-	-	-	-	-
4369290	+	0.0000	12	2	-	-	-	-	-	-	+
4371772	-	0.0000	68	5	-	-	-	-	-	-	+
4376389	-	0.0017	4	0	XCV3789	20	-	-	-	-	-
4377581	+	0.0009	11	3	XCV3791	25	-	-	XCV3790	83/976	-
4385950	+	0.0030	4	0	-	-	-	-	XCV3800	30/1145	-
4386005	-	0.0017	5	0	XCV3800	25	-	-	-	-	-
4386948	-	0.0021	4	0	-	-	-	-	XCV3801	10/270	-
4388305	-	0.0077	4	0	-	-	XCV3802	292/1042	-	-	-
4392400	+	0.0000	11	1	XCV3807	41	-	-	XCV3806	13/652	-
4393809	-	0.0194	4	2	-	-	XCV3808	538/556	-	-	-
4394950	-	0.0017	4	0	-	-	XCV3809	686/600	-	-	-
4395515	-	0.0030	4	0	-	-	XCV3809	121/1165	-	-	-
4398395	+	0.0001	8	0	XCV3814	26	-	-	-	-	-
4400663	+	0.0000	23	2	-	-	XCV3816	151/604	-	-	-
4402676	+	0.0018	9	3	-	-	XCV3820	49/301	-	-	-
4403069	+	0.0130	3	0	-	-	-	-	-	-	+
4407078	-	0.0020	3	0	XCV3823	68	-	-	-	-	-
4409801	-	0.0009	4	0	-	-	XCV3825	422/1011	-	-	-
4411740	-	0.0008	7	1	XCV3826	0	-	-	-	-	-
4413048	-	0.0400	4	0	XCV3827	252	-	-	XCV3828	85/978	-
4420847	+	0.0009	5	0	XCV3836	104	-	-	-	-	-
4423078	-	0.0000	12	0	-	-	XCV3837	245/1347	-	-	-
4423362	-	0.0192	6	1	XCV3837	39	-	-	-	-	-
4454565	+	0.0000	13	1	XCV3871	226	-	-	-	-	-
4480327	-	0.0021	6	0	XCV3894	262	XCV3895	199/232	-	-	-
4485330	-	0.0000	11	0	XCV3903	218	-	-	-	-	-
4488575	+	0.0019	3	0	-	-	-	-	-	-	+
4492811	-	0.0000	41	3	XCV3912	37	-	-	-	-	-
4492830	+	0.0125	3	0	-	-	-	-	XCV3913 XCV3912	493/56 56/1930	-
4493424	+	0.0109	4	1	-	-	XCV3914	102/911	-	-	-
4498709	-	0.0005	7	0	XCV3919	22	-	-	-	-	-
4498825	+	0.0000	16	0	-	-	-	-	-	-	+
4499967	+	0.0085	3	0	XCV3921	43	-	-	XCV3920	39/1013	-
4500051	-	0.0006	6	1	XCV3920	45	-	-	XCV3921	41/780	-
4507411	-	0.0004	7	1	XCV3926	42	-	-	XCV3927	90/1223	-
4507470	+	0.0000	19	5	XCV3927	31	-	-	-	-	-
4509798	+	0.0000	54	10	XCV3929	99	XCV3928	1089/53	-	-	-
4515928	+	0.0000	10	1	-	-	XCV3934	156/773	-	-	-
4518620	-	0.0000	168	23	XCV3936	297	-	-	XCV3937	38/250	-
4521105	-	0.0000	18	0	XCV3938	37	-	-	-	-	-

[a], TSS position on the *Xcv* chromosome and the plasmid (pXCV183; marked with #), respectively. [b], p-value indicates the significance of the annotated TSS (see METHODS). [c], number of read starts at TSS position. [d], TSS classification and the corresponding CDS. [e], distance of the TSS to annotated start and stop codon of the corresponding CDS. [f], a TSS is classified as orphan (+) if it is neither classified as primary, internal nor antisense TSS.

TSS position[a]	Strand	p-value[b]	Library 2[c]	Library 1[c]	Primary[d] to CDS	5' UTR (bp)	Internal[d] to CDS	Distance to start/stop[e] (bp)	Antisense[d] to CDS	Distance to start/stop[e] (bp)	Orphan[f]
4522609	+	0.0016	6	0	XCV3940	114	-	-	XCV3939	80/1453	-
4526218	+	0.0424	4	0	XCV3943	0	-	-	-	-	-
4533862	-	0.0000	86	11	-	-	XCV3949	9/872	-	-	-
4537266	-	0.0001	8	0	XCV3953	0	-	-	-	-	-
4539141	+	0.0000	16	2	XCV3955	121	-	-	XCV3954	78/1796	-
4539754	-	0.0167	4	0	-	-	-	-	-	-	+
4542837	-	0.0000	25	4	XCV3957	22	-	-	-	-	-
4548120	-	0.0002	8	0	-	-	XCV3960	101/2127	-	-	-
4550453	-	0.0002	7	1	XCV3963	0	-	-	-	-	-
4551938	-	0.0012	6	1	XCV3964	21	-	-	-	-	-
4562352	-	0.0000	10	0	-	-	XCV3969	75/596	-	-	-
4562806	+	0.0469	7	0	-	-	-	-	XCV3970	425/138	-
4563705	+	0.0177	3	1	-	-	-	-	XCV3971	422/327	-
4564127	-	0.0008	4	0	XCV3971	0	-	-	-	-	-
4566249	-	0.0006	6	0	XCV3976	71	-	-	-	-	-
4578759	-	0.0060	3	0	XCV3988	27	-	-	-	-	-
4580486	-	0.0047	3	0	XCV3990	23	-	-	-	-	-
4580654	+	0.0015	22	14	XCV3991	88	-	-	-	-	-
4583277	+	0.0041	6	2	XCV3994	0	-	-	-	-	-
4585144	+	0.0165	3	0	-	-	XCV3995	114/1034	-	-	-
4587148	+	0.0089	4	1	-	-	XCV3996	796/781	-	-	-
4589368	-	0.0000	8	0	-	-	XCV3998	862/298	-	-	-
4590816	-	0.0271	9	5	XCV3999	0	-	-	-	-	-
4596722	-	0.0000	22	4	-	-	XCV4007	147/500	-	-	-
4596940	+	0.0002	14	5	XCV4008	21	-	-	XCV4007	71/718	-
4606228	+	0.0261	5	0	-	-	-	-	-	-	+
4606630	-	0.0003	7	1	-	-	-	-	-	-	+
4608071	-	0.0000	11	2	XCV4010	33	-	-	-	-	-
4608171	+	0.0000	21	4	XCV4011	38	-	-	-	-	-
4615446	+	0.0000	12	0	XCV4018	30	-	-	-	-	-
4626132	+	0.0272	3	1	XCV4029	1	-	-	-	-	-
4627723	+	0.0051	9	4	XCV4030	299	XCV4029	1590/116	-	-	-
4628082	+	0.0415	5	2	-	-	XCV4030	60/821	-	-	-
4634680	+	0.0206	3	0	-	-	-	-	XCV4035	178/1606	-
4636053	+	0.0000	17	3	-	-	-	-	XCV4036	1018/1186	-
4641920	-	0.0001	8	0	XCV4040	31	-	-	-	-	-
4646308	-	0.0070	4	1	XCV4043	27	-	-	-	-	-
4647214	+	0.0010	6	0	XCV4046	73	-	-	-	-	-
4649419	+	0.0057	3	0	-	-	XCV4048	26/528	-	-	-
4652571	-	0.0001	4	0	-	-	XCV4052	907/211	-	-	-
4654564	-	0.0343	9	6	-	-	XCV4053	18/1073	XCV4054	93/737	-
4654612	+	0.0110	3	0	XCV4054	45	-	-	XCV4053	30/1121	-
4664567	+	0.0141	3	0	XCV4067	173	-	-	-	-	-
4665917	+	0.0024	3	0	-	-	XCV4068	143/519	-	-	-
4670222	-	0.0002	6	0	XCV4073	78	-	-	-	-	-
4674977	+	0.0001	13	3	XCV4079	0	-	-	-	-	-
4676672	+	0.0078	3	0	-	-	XCV4080	916/451	-	-	-
4691793	-	0.0012	4	0	XCV4094	27	-	-	XCV4095	20/523	-
4694340	-	0.0000	28	4	XCV4096	32	-	-	-	-	-
4694851	+	0.0043	3	0	-	-	-	-	XCV4097	880/139	-
4697095	-	0.0028	3	0	XCV4098	24	-	-	-	-	-
4698077	-	0.0000	105	8	XCV4100 XCV4101	268 41	-	-	-	-	-
4699089	-	0.0000	14	1	XCV4102	23	-	-	-	-	-

[a], TSS position on the *Xcv* chromosome and the plasmid (pXCV183; marked with #), respectively. [b], p-value indicates the significance of the annotated TSS (see METHODS). [c], number of read starts at TSS position. [d], TSS classification and the corresponding CDS. [e], distance of the TSS to annotated start and stop codon of the corresponding CDS. [f], a TSS is classified as orphan (+) if it is neither classified as primary, internal nor antisense TSS.

TSS position[a]	Strand	p-value[b]	Library 2[c]	Library 1[c]	Primary[d] to CDS	5' UTR (bp)	Internal[d] to CDS	Distance to start/stop[e] (bp)	Antisense[d] to CDS	Distance to start/stop[e] (bp)	Orphan[f]
4701994	-	0.0000	23	7	-	-	-	-	XCV4106	81/257	-
4706354	+	0.0070	4	1	-	-	XCV4111	10/301	-	-	-
4711085	+	0.0206	3	0	-	-	-	-	XCV4115	902/510	-
4714461	-	0.0015	4	0	XCV4118	44	-	-	-	-	-
4715131	-	0.0000	43	10	XCV4119	29	-	-	XCV4120	31/573	-
4720514	-	0.0088	4	0	XCV4123	1	-	-	-	-	-
4727508	-	0.0162	4	1	-	-	-	-	-	-	+
4728553	-	0.0000	15	1	XCV4131	20	-	-	-	-	-
4729028	-	0.0051	6	2	XCV4132	22	-	-	-	-	-
4731606	-	0.0104	3	0	-	-	XCV4134	17/1809	-	-	-
4735763	-	0.0011	7	2	XCV4137	177	-	-	-	-	-
4735903	+	0.0000	23	2	XCV4138	67	-	-	-	-	-
4736549	+	0.0008	6	0	-	-	XCV4138	579/2432	-	-	-
4737286	+	0.0422	3	0	-	-	XCV4138	1316/1695	-	-	-
4738388	+	0.0401	5	4	-	-	XCV4138	2418/593	-	-	-
4738712	+	0.0402	4	0	-	-	XCV4138	2742/269	-	-	-
4740424	+	0.0070	3	0	-	-	XCV4140	560/450	-	-	-
4748322	-	0.0132	5	1	XCV4147	160	-	-	-	-	-
4748727	-	0.0177	3	0	-	-	-	-	XCV4148	153/317	-
4751394	-	0.0007	4	0	-	-	XCV4151	23/801	XCV4152	60/1052	-
4751412	+	0.0006	7	1	XCV4152	42	-	-	XCV4151	5/819	-
4757260	+	0.0000	21	2	-	-	-	-	-	-	+
4757357	-	0.0000	12	3	-	-	-	-	-	-	+
4761503	-	0.0156	4	1	-	-	-	-	-	-	+
4768772	+	0.0000	14	1	XCV4166	17	-	-	-	-	-
4775972	-	0.0100	3	0	-	-	XCV4167	36/4274	-	-	-
4776441	-	0.0000	37	12	-	-	-	-	-	-	+
4784392	-	0.0001	13	3	XCV4170 XCV4171	287 34	-	-	-	-	-
4788659	-	0.0193	3	1	-	-	-	-	XCV4175	1294/325	-
4793007	-	0.0025	3	0	-	-	XCV4181	2017/346	-	-	-
4796313	+	0.0079	3	0	-	-	-	-	XCV4184 XCV4185	69/713 829/142	-
4798298	-	0.0023	7	2	XCV4187	111	XCV4188	689/54	-	-	-
4799322	-	0.0160	3	0	XCV4189	13	-	-	-	-	-
4804093	+	0.0039	3	0	XCV4194	26	-	-	-	-	-
4808258	-	0.0073	5	1	-	-	-	-	XCV4197	626/36	-
4808628	+	0.0000	17	1	XCV4198	0	-	-	-	-	-
4813172	-	0.0110	6	3	XCV4200	0	-	-	XCV4201	77/364	-
4857728	-	0.0087	3	0	XCV4227	23	-	-	-	-	-
4863179	+	0.0000	19	2	-	-	XCV4235	817/13	-	-	-
4882281	-	0.0001	7	0	-	-	XCV4251	86/1434	-	-	-
4887536	-	0.0000	15	2	XCV4257	162	-	-	-	-	-
4891644	+	0.0008	8	2	-	-	-	-	-	-	+
4895203	-	0.0000	14	2	XCV4262	0	-	-	-	-	-
4898418	-	0.0007	8	2	XCV4264	0	-	-	-	-	-
4902712	+	0.0151	3	0	XCV4268	41	-	-	XCV4267	53/679	-
4904300	-	0.0041	5	1	XCV4270	9	-	-	-	-	-
4906974	+	0.0031	4	0	-	-	-	-	XCV4273	40/274	-
4913268	+	0.0000	10	1	XCV4279	23	-	-	-	-	-
4914278	-	0.0017	6	1	-	-	-	-	-	-	+
4923729	-	0.0000	10	0	XCV4289	42	-	-	-	-	-
4946992	-	0.0000	9	0	XCV4306	0	-	-	-	-	-
4955269	-	0.0002	7	1	XCV4311	25	-	-	XCV4312	65/1876	-

[a], TSS position on the *Xcv* chromosome and the plasmid (pXCV183; marked with #), respectively. [b], p-value indicates the significance of the annotated TSS (see METHODS). [c], number of read starts at TSS position. [d], TSS classification and the corresponding CDS. [e], distance of the TSS to annotated start and stop codon of the corresponding CDS. [f], a TSS is classified as orphan (+) if it is neither classified as primary, internal nor antisense TSS.

TSS position[a]	Strand	p-value[b]	Library 2[c]	Library 1[c]	Primary[d] to CDS	5' UTR (bp)	Internal[d] to CDS	Distance to start/stop[e] (bp)	Antisense[d] to CDS	Distance to start/stop[e] (bp)	Orphan[f]
4966194	-	0.0004	8	2	-	-	XCV4315	42/1799	-	-	-
4968288	-	0.0126	3	0	-	-	-	-	XCV4318	845/84	-
4970190	-	0.0000	21	2	XCV4320 XCV4321	270 27	-	-	-	-	-
4971181	-	0.0000	20	0	XCV4322	37	-	-	-	-	-
4971307	+	0.0006	6	1	-	-	XCV4323	9/953	-	-	-
4973545	-	0.0041	4	0	-	-	-	-	-	-	+
4980145	-	0.0129	5	0	XCV4332	86	-	-	XCV4333	50/2283	-
4994114	-	0.0156	3	0	XCV4342	22	-	-	-	-	-
4994366	+	0.0134	3	0	XCV4343	22	-	-	-	-	-
5006692	-	0.0001	6	0	XCV4353	22	-	-	-	-	-
5010792	+	0.0013	13	0	-	-	-	-	XCV4356	93/2675	-
5012026	+	0.0135	5	0	-	-	-	-	XCV4357	362/1053	-
5019614	+	0.0206	3	0	XCV4363	39	-	-	-	-	-
5024943	+	0.0000	9	0	-	-	XCV4369	15/200	-	-	-
5031007	+	0.0000	17	4	XCV4373	137	-	-	-	-	-
5032259	+	0.0000	24	1	XCV4374	105	-	-	-	-	-
5036212	-	0.0257	3	0	-	-	-	-	XCV4375	493/637	-
5042497	-	0.0288	3	1	-	-	XCV4379	311/978	-	-	-
5045604	-	0.0003	5	0	XCV4382	48	-	-	-	-	-
5047832	+	0.0013	11	4	XCV4384	0	-	-	-	-	-
5047838	-	0.0096	3	0	XCV4383	212	-	-	XCV4384	6/620	-
5051823	-	0.0000	15	0	-	-	XCV4387	28/1621	-	-	-
5060407	+	0.0261	5	0	-	-	-	-	-	-	+
5060810	-	0.0086	5	1	-	-	-	-	-	-	+
5068657	-	0.0001	8	0	XCV4395	63	-	-	-	-	-
5070952	-	0.0002	6	0	-	-	XCV4399	3606/446	-	-	-
5084851	-	0.0027	3	0	XCV4408	27	-	-	-	-	-
5086838	+	0.0032	3	0	-	-	XCV4411	726/1118	-	-	-
5087567	-	0.0028	3	0	-	-	-	-	XCV4411	1455/389	-
5088118	+	0.0132	4	0	XCV4412	38	-	-	-	-	-
5094547	-	0.0008	6	0	-	-	-	-	-	-	+
5102472	+	0.0003	4	0	-	-	XCV4425	136/127	-	-	-
5103051	+	0.0028	3	0	XCV4426	18	-	-	-	-	-
5105859	+	0.0000	10	1	-	-	-	-	XCV4430	391/94	-
5106753	-	0.0061	4	1	-	-	XCV4431	457/16	-	-	-
5108588	+	0.0068	4	0	XCV4434	28	-	-	XCV4433	4/625	-
5110351	-	0.0466	4	0	-	-	XCV4435	195/626	-	-	-
5110576	-	0.0061	5	0	XCV4435	30	-	-	-	-	-
5110677	+	0.0008	4	0	XCV4436	130	-	-	-	-	-
5116206	-	0.0185	11	0	XCV4440	222	XCV4441	2272/100	-	-	-
5126147	-	0.0050	4	1	-	-	XCV4443	137/3414	-	-	-
5131412	-	0.0017	7	0	XCV4444	107	-	-	-	-	-
5134201	-	0.0000	10	1	XCV4449	25	-	-	-	-	-
5134535	+	0.0000	17	5	XCV4450	24	-	-	-	-	-
5139920	-	0.0007	5	0	XCV4457	21	-	-	-	-	-
5142226	-	0.0414	10	0	XCV4459	138	-	-	-	-	-
5146436	+	0.0002	6	0	XCV4466	26	-	-	-	-	-
5153717	-	0.0193	3	0	XCV4468	112	-	-	-	-	-
5154683	-	0.0047	3	0	XCV4469	22	-	-	-	-	-
5156544	-	0.0002	8	0	-	-	XCV4470	25/1768	-	-	-
5164171	-	0.0166	3	1	XCV4476	0	-	-	-	-	-
5164230	+	0.0072	3	0	XCV4477	206	-	-	XCV4476	59/1222	-
5165820	+	0.0130	4	1	XCV4478	0	-	-	-	-	-

[a], TSS position on the *Xcv* chromosome and the plasmid (pXCV183; marked with #), respectively. [b], p-value indicates the significance of the annotated TSS (see METHODS). [c], number of read starts at TSS position. [d], TSS classification and the corresponding CDS. [e], distance of the TSS to annotated start and stop codon of the corresponding CDS. [f], a TSS is classified as orphan (+) if it is neither classified as primary, internal nor antisense TSS.

TSS position[a]	Strand	p-value[b]	Library 2[c]	Library 1[c]	Primary[d] to CDS	5' UTR (bp)	Internal[d] to CDS	Distance to start/stop[e] (bp)	Antisense[d] to CDS	Distance to start/stop[e] (bp)	Orphan[f]
5170524	+	0.0000	14	0	*XCV4482*	28	-	-	-	-	-
5178255	-	0.0000	293	20	*XCV4486* *XCV4487*	210 31	-	-	-	-	-
5178392	+	0.0000	23	2	-	-	-	-	-	-	+

Reference

1. Thieme, F., Koebnik, R., Bekel, T., Berger, C., Boch, J., Büttner, D., Caldana, C., Gaigalat, L., Goesmann, A., Kay, S. *et al.* (2005) Insights into genome plasticity and pathogenicity of the plant pathogenic bacterium *Xanthomonas campestris* pv. vesicatoria revealed by the complete genome sequence. *J. Bacteriol.*, **187**, 7254-7266.

[a], TSS position on the *Xcv* chromosome and the plasmid (pXCV183; marked with #), respectively. [b], p-value indicates the significance of the annotated TSS (see METHODS). [c], number of read starts at TSS position. [d], TSS classification and the corresponding CDS. [e], distance of the TSS to annotated start and stop codon of the corresponding CDS. [f], a TSS is classified as orphan (+) if it is neither classified as primary, internal nor antisense TSS.

Table S3. Putative chromosomal TSSs within the first 50 bp of annotated CDSs (1).

CDS	Strand	Annotated translation start	TSS position	Library 2[a]	Library 1[a]	Distance of annotated start codon and TSS (bp)
XCV0016	+	17004	17027	42	18	24
XCV0022	+	23353	23398	13	1	46
XCV0035	-	38772	38760	13	0	13
XCV0117	+	133981	133989	3	1	9
XCV0209	+	244583	244587	15	2	5
XCV0322	+	369633	369663	3	0	31
XCV0327	-	375518	375494	3	0	25
XCV0383	+	436658	436669	4	0	12
XCV0387	-	441649	441623	4	0	27
XCV0493	+	557939	557964	3	0	26
XCV0522	-	589469	589427	10	0	43
XCV0537	+	602285	602321	16	1	37
XCV0579	+	656595	656637	6	2	43
XCV0632	-	713619	713610	3	0	10
XCV0648	+	732783	732826	4	0	44
XCV0653	-	736984	736976	3	0	9
XCV0781	+	893350	893392	6	0	43
XCV0816	-	931162	931121	3	0	42
XCV0821	-	936169	936154	12	1	16
XCV1073	-	1206623	1206581	3	0	43
XCV1074	+	1206680	1206687	11	1	8
XCV1141	+	1270899	1270936	47	4	38
XCV1155	-	1284633	1284606	4	1	28
XCV1239	-	1391023	1390988	3	0	36
XCV1332	-	1498481	1498473	8	0	9
XCV1400	-	1581081	1581048	5	1	34
XCV1465	-	1648439	1648431	5	0	9
XCV1499	+	1690119	1690128	3	0	10
XCV1586	-	1793475	1793470	7	0	6
XCV1611	-	1825244	1825238	8	1	7
XCV1617	-	1830765	1830728	3	0	38
XCV1654	+	1864564	1864572	7	0	9
XCV1922	+	2173724	2173764	3	1	41
XCV2017	-	2299645	2299637	29	3	9
XCV2073	+	2363433	2363454	8	0	22
XCV2207	+	2536959	2536986	11	2	28
XCV2320	+	2651486	2651497	151	12	12
XCV2329	-	2660238	2660227	151	14	12
XCV2340	-	2672433	2672422	151	14	12
XCV2341	-	2673253	2672422	151	14	12
XCV2353	+	2682750	2682761	151	12	12
XCV2482	+	2801007	2801025	5	0	19
XCV2535	+	2859384	2859395	5	1	12
XCV2729	-	3103384	3103362	3	0	23
XCV2750	+	3130305	3130321	5	0	17
XCV2770	-	3151135	3151089	4	0	47
XCV2971	-	3379738	3379701	7	1	38
XCV2972	-	3380697	3380658	4	0	40
XCV2974	-	3385412	3385409	12	1	4
XCV3037	-	3459123	3459115	10	5	9
XCV3065	+	3488846	3488858	8	0	13
XCV3123	+	3539930	3539954	11	0	25
XCV3188	-	3624084	3624042	14	2	43
XCV3257	+	3725390	3725397	7	0	8
XCV3392	-	3880185	3880139	52	3	47
XCV3491	-	3999770	3999749	3	0	22
XCV3532	+	4042171	4042199	3	0	29
XCV3625	-	4171917	4171902	5	1	16

[a], number of read starts at TSS position.

CDS	Strand	Annotated translation start	TSS position	Library 2[a]	Library 1[a]	Distance of annotated start codon and TSS (bp)
XCV3746	+	4328075	4328091	13	3	17
XCV3949	-	4533871	4533862	86	11	10
XCV4048	+	4649393	4649419	3	0	27
XCV4053	-	4654582	4654564	9	6	19
XCV4111	+	4706344	4706354	4	1	11
XCV4134	-	4731623	4731606	3	0	18
XCV4151	-	4751417	4751394	4	0	24
XCV4167	-	4776008	4775972	3	0	37
XCV4315	-	4966236	4966194	8	2	43
XCV4323	+	4971298	4971307	6	1	10
XCV4369	+	5024928	5024943	9	0	16
XCV4387	-	5051851	5051823	15	0	29
XCV4470	-	5156569	5156544	8	0	26

Reference

1. Thieme, F., Koebnik, R., Bekel, T., Berger, C., Boch, J., Büttner, D., Caldana, C., Gaigalat, L., Goesmann, A., Kay, S. *et al.* (2005) Insights into genome plasticity and pathogenicity of the plant pathogenic bacterium *Xanthomonas campestris* pv. vesicatoria revealed by the complete genome sequence. *J. Bacteriol.*, **187**, 7254-7266.

[a], number of read starts at TSS position.

Table S4. Predicted antisense TSSs located close to the 3' end of annotated CDSs (+/- 100 bp) (1).

TSS position[a]	Strand	Library 2[b]	Library 1[b]	Distance of CDS end to TSS[c] (bp)	Gene product
81749#	-	9	2	43	XCVd0073: hypothetical protein
109851#	+	35	6	-99	XCVd0099: putative zeta toxin of the postsegregational killing system
127227#	-	4	2	-62	XCVd0115: Tn5044 transposase
159633#	+	3	0	-74	XCVd0154: hypothetical protein
180140#	+	3	0	-86	XCVd0171: hypothetical protein
11675	-	6	0	-72	XCV0008: energy transducer TonB protein
78522	+	23	2	-46	XCV0067: hypothetical protein
204493	+	12	2	-71	XCV0172: indolepyruvate ferredoxin oxidoreductase
261741	-	15	1	3	XCV0221: hypothetical protein
321787	-	7	0	-77	XCV0275: putative DNA-binding protein
485143	-	13	0	-94	XCV0435: HrcC protein
533893	+	6	0	-36	XCV0472: hypothetical protein
558968	-	8	1	-82	XCV0493: hypothetical protein
590048	-	10	0	-13	XCV0523: 50S ribosomal protein L13
707695	+	4	1	-8	XCV0625: IS1477 transposase
799920	-	6	2	-49	XCV0704: hypothetical protein
1041488	-	3	1	21	XCV0911: putative secreted protein
1095646	-	4	0	-86	XCV0964: hypothetical protein
1454260	-	3	1	-15	XCV1295: putative phage replication protein
1618915	+	4	0	-35	XCV1441: methionyl-tRNA synthetase
1659403	+	4	0	93	XCV1477: putative signal transduction protein
1750856	+	21	2	-46	XCV1552: hypothetical protein
1783909	+	5	0	-52	XCV1579: hypothetical protein
1875752	-	6	0	-78	XCV1666: hypothetical protein
1880820	-	3	0	-42	XCV1669: lysyl-tRNA synthetase
2078061	+	4	1	-8	XCV1840: IS1477 transposase
2081955	+	3	0	-76	XCV1846: ISxcC1 transposase
2191418	+	4	0	54	XCV1936: hypothetical protein
2192469	+	3	0	20	XCV1937: hypothetical protein
2484379	-	210	29	-75	XCV2161: conjugal transfer protein TrbL
2580164	+	3	0	-29	XCV2250: putative secreted protein
2595113	-	5	1	-68	XCV2261: phage-related integrase
2656549	+	8	4	-35	XCV2325: streptomycin 6-kinase
2768989	-	10	3	-67	XCV2436: IS1595 transposase
2799097	+	5	1	-68	XCV2479: putative replication protein C (fragment)
2813614	+	3	0	-90	XCV2492: glycosyltransferase
2885613	-	6	1	43	XCV2557: hypothetical protein
3151886	+	4	1	-8	XCV2772: IS1477 transposase
3206053	+	4	1	-8	XCV2816: IS1477 transposase
3845146	-	3	1	84	XCV3357: putative secreted protein
3873924	+	20	2	-46	XCV3388: hypothetical protein
3899941	-	5	0	-64	XCV3412: ISxac2 transposase
4027759	-	3	0	-55	XCV3514: ribonuclease PH
4367233	-	11	0	-77	XCV3782: OmpW family outer membrane protein
4367880	-	8	1	47	XCV3783: putative secreted protein
4492830	+	3	0	-56	XCV3913: D-tyrosyl-tRNA(Tyr) deacylase
4808258	-	5	1	-36	XCV4197: hypothetical protein
4968288	-	3	0	84	XCV4318: hypothetical protein
5105859	+	10	1	94	XCV4430: hypothetical protein

[a], TSS position on the *Xcv* chromosome and the plasmid (pXCV183; marked with #), respectively. [b], number of read starts at TSS position. [c], positive and negative sign indicates TSS within and downstream the CDS, respectively.

Reference

1. Thieme, F., Koebnik, R., Bekel, T., Berger, C., Boch, J., Büttner, D., Caldana, C., Gaigalat, L., Goesmann, A., Kay, S. *et al.* (2005) Insights into genome plasticity and pathogenicity of the plant pathogenic bacterium *Xanthomonas campestris* pv. vesicatoria revealed by the complete genome sequence. *J. Bacteriol.*, **187**, 7254-7266.

Table S5. 5' UTRs of type III effector genes, leaderless genes with TSSs ≤10 bp upstream of the annotated start codon and genes containing long 5' UTRs (150 to 300 bp) referring to the annotated genome sequence of *Xcv* strain 85-10 (1). TSSs marked with # are located on pXCV183. TSSs of type III effector genes marked in bold exhibit a PIP box and -10 T/A-rich element as previously described (2). The PIP box of the type III effector gene *XCV2280* was identified in this study and is shown in brackets.

TSS position	Strand	Library 2[a]	Library 1[a]	5' UTR (bp)	Gene product
5' UTRs of Type III Effectors					
114827#	-	11	0	375	XCVd0104: avirulence protein AvrBs1
62304	+	3	0	23	XCV0052: avirulence protein AvrBs2
486511	+	5	0	678	XCV0437: xanthomonas outer protein D (XopD)
532770	+	8	1	0	XCV0471: avirulence protein AvrRxv
660302	+	12	2	341	XCV0581: xanthomonas outer protein B (XopB)
1184957	+	12	1	25	XCV1055: xanthomonas outer protein O (XopO)
2480236	-	183	10	22	XCV2156: xanthomonas outer protein J1 (XopJ1)
2613410 (2613343-2613403)	+	33	2	23	XCV2280: xanthomonas outer protein E2 (XopE2)
2767760	-	8	1	187	XCV2435: xanthomonas outer protein C (XopC)
3348398	+	12	3	173	XCV2944: xanthomonas outer protein N (XopN)
4371772	-	68	5	477	XCV3785: xanthomonas outer protein AA (XopAA)
Leaderless Transcripts					
193010	+	9	2	0	XCV0161: putative amidase/aminoacylase/peptidase family protein
201827	-	7	2	0	XCV0168: putative secreted protein
238550	-	7	0	0	XCV0201: uroporphyrinogen-III synthase
278015	-	6	0	0	XCV0236: predicted rRNA methylase
283462	+	4	0	0	XCV0243: acyl-CoA synthetase
288111	+	10	5	0	XCV0247: hypothetical protein
332993	+	9	1	0	XCV0287: predicted hydrolase of the alpha/beta superfamily
377721	+	4	0	3	XCV0330: putative NADH flavin oxidoreductase
411941	+	3	1	6	XCV0359: predicted Fe-S-cluster oxidoreductase
532770	+	8	1	0	XCV0471: avirulence protein AvrRxv
585332	+	8	5	0	XCV0518: phosphoserine phosphatase
599405	+	18	5	0	XCV0534: putative Iron-sulfur cluster assembly accessory protein
770365	+	4	0	0	XCV0679: porphobilinogen deaminase
779535	-	8	3	0	XCV0686: aspartate/tyrosine/aromatic aminotransferase
782120	+	7	0	0	XCV0688: hypothetical protein
823961	+	4	0	0	XCV0723: membrane-bound lytic transglycosylase
851592	+	8	0	0	XCV0747: 2-amino-4-hydroxy-6-hydroxymethyldihydropteridine pyrophosphokinase
884284	-	4	0	0	XCV0773: tRNA nucleotidyltransferase/poly(A) polymerase
959187	+	9	2	0	XCV0840: thiamine monophosphate synthase
997844	-	21	0	0	XCV0868: tRNA-dihydrouridine synthase
1028253	-	3	0	0	XCV0899: hypothetical protein
1029078	-	5	0	0	XCV0900: dimethyladenosine transferase (rRNA methylation)
1036560	+	6	0	0	XCV0907: 2-octaprenyl-6-methoxyphenyl hydroxylase
1102907	-	3	1	0	XCV0973: molybdenum cofactor biosynthesis protein B
1265655	-	3	1	0	XCV1134: MoxR-like ATPase
1419087	+	8	1	0	XCV1260: two-component system sensor protein
1450021	+	8	1	0	XCV1292: 4-hydroxy-3-methylbut-2-enyl diphosphate reductase
1504286	-	15	0	0	XCV1337: putative cytochrome C assembly protein
1601206	+	5	2	0	XCV1423: hypothetical protein
1631636	-	3	0	8	XCV1452: hypothetical protein
1631992	+	3	1	10	XCV1453: hypothetical protein
1669473	+	12	1	0	XCV1485: methionine aminopeptidase
1712207	-	4	0	0	XCV1516: ABC transporter permease and ATP-binding protein
1719682	+	20	2	0	XCV1525: NAD/FAD-binding protein
1756686	-	27	2	0	XCV1556: tmRNA-binding protein SmpB
1816736	+	17	0	1	XCV1607: RNase R

[a], number of read starts at the TSS position.

TSS position	Strand	Library 2[a]	Library 1[a]	5' UTR (bp)	Gene product
1843191	+	5	0	0	XCV1631: hypothetical protein
1859472	-	7	2	0	XCV1646: carbonic anhydrase
1988882	+	9	2	9	XCV1768: RNA-binding protein Hfq
1993650	+	4	1	0	XCV1772: LexA repressor
2046910	+	3	0	10	XCV1814: tRNA nucleotidyltransferase/poly(A) polymerase
2056210	+	3	0	3	XCV1823: Glucan 1,4-beta-glucosidase precursor
2066165	-	6	3	0	XCV1828: two-component system sensor protein
2066489	+	7	0	0	XCV1829: 4-hydroxy-3-methylbut-2-en-1-yl diphosphate synthase
2133276	-	11	2	0	XCV1886: methylthioribulose-1-phosphate dehydratase
2316851	-	7	0	0	XCV2034: chemotaxis signal transduction protein
2349615	+	5	0	0	XCV2060: putative ATPase
2363433	+	8	0	0	XCV2073: molybdopterin biosynthesis protein
2411526	-	3	0	0	XCV2105: two-component system sensor kinase
2534689	+	11	4	0	XCV2204: hypothetical protein
2646914	+	3	0	1	XCV2314: ArsR family transcriptional regulator
2744537	-	3	0	9	XCV2419: FKBP-type peptidyl-prolyl cis-trans isomerase
2744576	+	7	4	0	XCV2420: putative membrane protein
2745041	+	4	0	0	XCV2421: predicted sulfurtransferase
2756332	+	14	3	0	XCV2429: thioredoxin-like protein
2843374	+	8	0	0	XCV2519: glutamine cyclotransferase
2853818	+	8	1	0	XCV2531: homoserine O-acetyltransferase
2915414	-	21	2	0	XCV2586: ABC transporter ATPase
2915525	+	5	1	0	XCV2587: ATP-dependent RNA helicase
2927298	+	3	0	4	XCV2597: glutamyl-tRNA synthetase
2935603	-	17	4	0	XCV2603: N-acetylmuramoyl-L-alanine amidase precursor
2937657	-	4	1	0	XCV2605: hypothetical protein
2966234	-	8	1	0	XCV2630: stringent starvation protein B
2971502	-	31	4	0	XCV2635: putative soluble lytic murein transglycosylase precursor
3024425	+	3	0	0	XCV2675: TetR family transcriptional regulator
3065332	-	14	0	0	XCV2704: cellulase precursor
3129258	-	71	9	0	XCV2748: proline racemase
3142239	+	3	0	4	XCV2763: HNH endonuclease family protein
3304605	-	3	0	0	XCV2902: putative 6-O-methylguanine-DNA methyltransferase
3317782	+	48	8	0	XCV2915: exodeoxyribonuclease VII small subunit
3325124	-	8	3	0	XCV2923: putative modulator of DNA gyrase
3334516	-	17	1	0	XCV2931: rRNA large subunit methyltransferase
3339362	-	14	0	0	XCV2936: rare lipoprotein B
3377499	+	8	0	0	XCV2970: hypothetical protein
3390399	+	3	0	0	XCV2983: 3-methyladenine DNA glycosylase
3410761	+	15	3	6	XCV3000: LysR family transcriptional regulator
3429740	+	4	0	0	XCV3016: hypothetical protein
3450047	+	4	1	0	XCV3030: deoxyribonuclease V
3487971	-	7	0	0	XCV3063: hypothetical protein
3505699	+	18	4	0	XCV3083: acetyltransferase (GNAT) family protein
3518387	-	994	98	0	XCV3096: ComEA-related DNA uptake protein
3533944	+	4	1	2	XCV3116: putative secreted protein
3539930	+	24	15	0	XCV3123: putative enzyme II of the phosphotransferase system
3583485	+	5	1	0	XCV3162: hypothetical protein
3601849	+	8	3	0	XCV3169: hypothetical protein
3639271	+	16	2	0	XCV3202: adenosylmethionine--8-amino-7-oxononanoate transaminase
3744037	-	6	1	0	XCV3277: Holliday junction DNA helicase
3862413	+	7	3	0	XCV3376: hypothetical protein
3903440	+	13	1	0	XCV3416: thiamin S sulphur transfer protein
3963400	+	26	4	0	XCV3461: hypothetical protein
3991134	+	3	0	0	XCV3486: putative secreted protein
4007577	-	28	2	0	XCV3496: MoxR-like ATPase
4024561	-	9	1	0	XCV3510: guanosine-3',5'-bis(diphosphate) 3'-pyrophosphohydrolase

[a], number of read starts at the TSS position.

TSS position	Strand	Library 2[a]	Library 1[a]	5' UTR (bp)	Gene product
4064199	+	5	1	0	XCV3548: putative membrane-associated phospholipid phosphatase
4069816	-	7	1	0	XCV3552: UDP-N-acetylmuramate--L-alanine ligase
4122387	+	8	2	0	XCV3589: TetR family transcriptional regulator
4199423	-	6	0	2	XCV3647: nicotinate phosphoribosyltransferase
4264986	-	16	5	0	XCV3690: CDP-diacylglycerol--serine O-phosphatidyltransferase
4318420	+	7	1	0	XCV3737: putative signal transduction protein
4411740	-	7	1	0	XCV3826: hypothetical protein
4526218	+	4	0	0	XCV3943: primosomal protein N
4537266	-	8	0	0	XCV3953: ABC transporter ATP-binding protein involved in cell division
4550453	-	7	1	0	XCV3963: hypothetical protein
4564127	-	4	0	0	XCV3971: HAD superfamily hydrolase
4583277	+	6	2	0	XCV3994: DNA primase
4590816	-	9	5	0	XCV3999: hypothetical protein
4626132	+	3	1	1	XCV4029: arginyl-tRNA synthetase
4674977	+	13	3	0	XCV4079: predicted hydrolase (HAD superfamily)
4720514	-	4	0	1	XCV4123: ATP-dependent DNA helicase
4808628	+	17	1	0	XCV4198: coproporphyrinogen III oxidase
4813172	-	6	3	0	XCV4200: DNA polymerase I
4895203	-	14	2	0	XCV4262: putative secreted protein
4898418	-	8	2	0	XCV4264: methyltransferase
4904300	-	5	1	9	XCV4270: exodeoxyribonuclease III
4946992	-	9	0	0	XCV4306: putative outer membrane protein
5047832	+	11	4	0	XCV4384: thymidine kinase
5164171	-	3	1	0	XCV4476: hypothetical protein
5165820	+	4	1	0	XCV4478: putative membrane protein
long 5' UTRs					
101600#	+	5	0	287	XCVd0093: putative secreted protein
116378#	+	80	8	277	XCVd0107: ISxac2 transposase (fragment)
118654#	+	8	1	201	XCVd0110: hypothetical protein
123136#	+	5	0	218	XCVd0114: putative cointegrate resolution protein S
1404	+	15	1	202	XCV0002: DNA polymerase III subunit beta chain
78978	-	8	5	225	XCV0067: hypothetical protein
81973	-	6	2	270	XCV0069: putative secreted protein
117077	-	4	0	274	XCV0099: putative secreted protein
155750	+	11	2	227	XCV0131: hypothetical protein
253127	-	25	10	203	XCV0214: putative transposase
304274	+	5	0	150	XCV0264: malate synthase
320398	+	6	1	234	XCV0274: TetR family transcriptional regulator
448950	+	3	0	226	XCV0395: hypothetical protein
532651	-	5	0	151	XCV0470: ATP-dependent RNA helicase
647837	+	18	0	192	XCV0574: phospho-2-dehydro-3-deoxyheptonate aldolase
765474	+	3	0	282	XCV0675: glucans biosynthesis glucosyltransferase H
765583	+	3	0	173	XCV0675: glucans biosynthesis glucosyltransferase H
782575	+	13	2	244	XCV0689: diaminopimelate epimerase
829714	+	24	4	266	XCV0728: lipoate-protein ligase B
951830	+	10	2	181	XCV0835: cell division protein FtsZ
1042407	+	7	1	180	XCV0913: protocatechuate 4,5-dioxygenase beta chain
1116481	+	7	2	259	XCV0986: transcription antitermination factor
1137159	+	79	14	180	XCV1002: 50S ribosomal protein L2
1137964	+	9	2	209	XCV1003: 30S ribosomal protein S19
1149059	+	4	1	233	XCV1024: 50S ribosomal protein L17
1149619	+	3	0	180	XCV1025: hypothetical protein
1164003	-	5	0	163	XCV1036: hypothetical protein
1190312	+	8	2	228	XCV1061: Glutamine phosphoribosylpyrophosphate amidotransferase
1250212	+	5	1	153	XCV1120: hypothetical protein
1340207	+	6	1	233	XCV1198: hypothetical protein
1399557	-	4	2	291	XCV1243: glycine cleavage system P-protein

[a], number of read starts at the TSS position.

TSS position	Strand	Library 2[a]	Library 1[a]	5' UTR (bp)	Gene product
1424474	-	4	0	171	XCV1265: putative D-alanyl-D-alanine carboxypeptidase
1549864	+	6	1	189	XCV1370: RNA polymerase ECF-type sigma factor RpoE2
1555991	+	7	2	180	XCV1375: putative membrane protein
1580759	-	6	1	248	XCV1399: putative secreted protein
1591652	-	5	1	154	XCV1410: hypothetical protein
1653692	-	18	6	209	XCV1470: outer membrane antigen
1663194	-	15	4	242	XCV1479: 30S ribosomal protein S2
1676072	+	7	0	220	XCV1490: asparagine synthetase B
1797706	-	9	0	280	XCV1591: GntR family transcriptional regulator
1998924	+	18	1	249	XCV1776: carbon storage regulator
2011844	-	4	0	264	XCV1789: hypothetical protein
2081259	+	5	1	156	XCV1845: putative membrane protein
2109973	+	20	2	151	XCV1864: hypothetical protein
2111307	+	17	4	226	XCV1866: bifunctional aspartokinase/homoserine dehydrogenaseI
2136615	-	4	0	270	XCV1888: amino acid-polyamine-organocation superfamily protein
2236742	-	5	0	157	XCV1962: Flagellar motor component MotA
2339660	+	10	1	270	XCV2056: cell division protein FtsK
2350997	+	5	2	171	XCV2061: Putative membrane-associated chromosome condensation protein
2399163	+	17	2	185	XCV2103: putative filamentous hemagglutinin-like protein
2630585	+	5	1	207	XCV2297: ISxac2 transposase
2631738	+	5	1	211	XCV2298: phage-related integrase
2676073	+	17	4	205	XCV2345: putative cytosine-specific DNA methylase
2767760	-	8	1	187	XCV2435: xanthomonas outer protein C (XopC)
2803838	-	24	7	162	XCV2484: phage-related integrase
2811009	-	3	1	296	XCV2489: glutathione-regulated potassium efflux protein B
2833387	-	5	1	242	XCV2509: putative secreted protein
2880224	-	36	5	264	XCV2551: N-acetylornithine carbamoyltransferase
2894891	+	45	4	229	XCV2568: putative secreted protein
2894892	+	30	0	228	XCV2568: putative secreted protein
2921819	-	349	6	157	XCV2592: major cold shock protein
2953268	+	16	2	279	XCV2620: putative DNA-binding protein
2959110	+	3	0	290	XCV2626: 3-methyladenine DNA glycosylase
3011678	+	3	0	212	XCV2671: two-component system sensor histidine kinase-response regulator hybridprotein
3045902	-	5	1	242	XCV2688: hypothetical protein
3094539	-	8	0	242	XCV2725: TonB-dependent outer membrane receptor
3106814	-	16	1	214	XCV2730: TonB-dependent outer membrane receptor
3174731	-	7	1	263	XCV2794: translation initiation factor IF-3
3196040	-	12	1	257	XCV2809: putative secreted protein
3247561	-	10	1	185	XCV2854: preprotein translocase subunit SecG
3259835	-	5	1	289	XCV2865: tryptophan synthase subunit alpha
3348398	+	12	3	173	XCV2944: xanthomonas outer protein N (XopN)
3426246	+	10	2	178	XCV3013: putative cysteine protease precursor
3478034	-	6	0	173	XCV3053: UDP-N-acetylmuramoyl-L-alanyl-D-glutamate synthetase
3496288	+	13	3	199	XCV3072: membrane-bound metalloendopeptidase
3531766	-	12	1	281	XCV3111: hypothetical protein
3610521	-	53	4	188	XCV3177: homoserine O-acetyltransferase
3748830	-	3	0	265	XCV3281: Holliday junction resolvase
3819837	+	21	0	152	XCV3337: glucose dehydrogenase
3832104	-	5	0	170	XCV3345: NAD synthetase
3872641	-	20	0	179	XCV3384: ISxac3 transposase
3877920	-	26	2	247	XCV3391: hypothetical protein
3886210	+	17	6	173	XCV3399: hypothetical protein
3969443	-	3	0	258	XCV3465: phosphoglycerate kinase
4012132	-	6	0	174	XCV3500: Tfp pilus assembly protein PilN
4037253	+	14	7	218	XCV3525: hypothetical protein
4100738	-	25	1	184	XCV3575: thiamine biosynthesis protein ThiC

[a], number of read starts at the TSS position.

TSS position	Strand	Library 2[a]	Library 1[a]	5' UTR (bp)	Gene product
4112499	+	3	0	186	XCV3583: 2-isopropylmalate synthase
4258733	+	4	0	283	XCV3686: soluble lytic murein transglycosylase precursor
4270597	+	4	0	248	XCV3696: ABC transporter ATP-binding protein
4282216	-	7	0	260	XCV3705: UDP-glucose dehydrogenase
4309871	-	9	3	186	XCV3727: hypothetical protein
4314757	+	7	2	241	XCV3733: ATP-dependent RNA helicase
4326863	-	3	1	182	XCV3744: oligopeptidase A
4413048	-	4	0	252	XCV3827: hypothetical protein
4454565	+	13	1	226	XCV3871: hypothetical protein
4480327	-	6	0	262	XCV3894: methionine aminopeptidase
4485330	-	11	0	218	XCV3903: hypothetical protein
4518620	-	168	23	297	XCV3936: tropinone reductase
4627723	+	9	4	299	XCV4030: putative secreted protein
4664567	+	3	0	173	XCV4067: hypothetical protein
4735763	-	7	2	177	XCV4137: putative glutathione transferase
4748322	-	5	1	160	XCV4147: putative membrane protein
4887536	-	15	2	162	XCV4257: 50S ribosomal protein L28
5047838	-	3	0	212	XCV4383: ATP-dependent DNA helicase
5116206	-	11	0	222	XCV4440: hypothetical protein
5164230	+	3	0	206	XCV4477: tellurium resistance protein

References

1. Thieme, F., Koebnik, R., Bekel, T., Berger, C., Boch, J., Büttner, D., Caldana, C., Gaigalat, L., Goesmann, A., Kay, S. *et al.* (2005) Insights into genome plasticity and pathogenicity of the plant pathogenic bacterium *Xanthomonas campestris* pv. vesicatoria revealed by the complete genome sequence. *J. Bacteriol.*, **187**, 7254-7266.
2. Koebnik, R., Krüger, A., Thieme, F., Urban, A. and Bonas, U. (2006) Specific binding of the *Xanthomonas campestris* pv. vesicatoria AraC-type transcriptional activator HrpX to plant-inducible promoter boxes. *J. Bacteriol.*, **188**, 7652-7660.

[a], number of read starts at the TSS position.

Table S6. Candidate riboswitches and widely conserved RNAs in *Xcv*.

Riboswitch (Rfam entry)[a]	Chromosomal position (strand)[b]	TSS position[c]	Downstream CDS[d]
FMN (RF00050)	911608-911787 (+)	-	Riboflavin synthase subunit alpha (XCV0800)
SAH (RF01057)	976910-977023 (-)	977026 (5/0)	S-adenosyl-L-homocysteine hydrolase (XCV0856)
Glycine (RF00504)	1399365-1399480 (-)	1399557 (4/2)	Glycine dehydrogenase (XCV1243)
	1399482-1399550 (-)	-	Glycine dehydrogenase (XCV1243)
SAM (RF00162)	3610402-3610523 (-)	3610521 (53/4)	Homoserine O-acetyltransferase (XCV3177)
Cobalamin (RF00174)	3801922-3802147 (-)	-	TonB-dependent outer membrane receptor (XCV3323)
TPP (RF00059)	4100645-4100743 (-)	4100738 (25/1)	Thiamine biosynthesis protein ThiC (XCV3575)
YybP-ykoY candidate (RF00080)	4924876-4924995 (+)	-	Hypothetical/putative membrane protein (XCV4291)
YybP-ykoY candidate (RF00080)	5164221-5164342 (+)	5164230 (3/0)	Tellurium resistance export protein (XCV4477)

sRNA (Rfam entry)[a]	Chromosomal position (strand)[b]	TSS position[c]	Chromosomal context[e]
RNaseP bacterial (RF00010)	933945-934295 (+)	933943 (86/27)	--> XCV0818 --> RNaseP --> XCV819
RtT (RF00391)	1235501-1235623 (+)	-	--> XCV1100 --> RtT --> XCV1101
SRP bacterial (RF00169)	1241912-1242175 (-)	1242175 (91/11)	--> XCV1109 <-- SRP <-- XCV1110
tmRNA (RF00023)	1753051-1753447 (-)	1753470 (31/1)	<-- XCV1553 <-- tmRNA --> XCV1554
6S RNA (RF00013)	4037865-4038049 (+)	4037865 (2627/442)	--> XCV3525 --> 6S RNA --> XCV3526

[a], Putative riboswitches and widely conserved RNAs identified in *Rfam* database [see SI; (1)]. [b], chromosomal position in *Xcv* corresponding to *Rfam* entry. [c], TSS position in the *Xcv* chromosome (2). Numbers in brackets indicate read starts at TSS position (library 2/library 1). [d], gene product of CDS annotated downstream of putative riboswitch (2). [e], chromosomal orientation of sRNA gene and flanking CDSs indicated by arrows.

References

1. Gardner, P.P., Daub, J., Tate, J., Moore, B.L., Osuch, I.H., Griffiths-Jones, S., Finn, R.D., Nawrocki, E.P., Kolbe, D.L., Eddy, S.R. *et al.* (2010) Rfam: Wikipedia, clans and the "decimal" release. *Nucleic Acids Res.*, **39**, D141-145.
2. Thieme, F., Koebnik, R., Bekel, T., Berger, C., Boch, J., Büttner, D., Caldana, C., Gaigalat, L., Goesmann, A., Kay, S. *et al.* (2005) Insights into genome plasticity and pathogenicity of the plant pathogenic bacterium *Xanthomonas campestris* pv. vesicatoria revealed by the complete genome sequence. *J. Bacteriol.*, **187**, 7254-7266.

Table S7. RNAcode prediction of novel chromosomal ORFs in *Xcv* (1, 2). Alternative start codons are listed under "chromosomal position". RNAcode6 is highlighted in bold and corresponds to sX6 which was experimentally verified (see Figure 3).

RNAcode prediction	Chromosomal position (Strand)	Transcript Support[a]	TSS position[b]
RNAcode1	285154-285243 (+)	0/0	-
RNAcode2	465113-465232 (+)	0/0	-
RNAcode3	722692-722736 (+)	1/0	-
RNAcode4	1281664-1281846 (+) 1281757-1281846 (+)	1/1	-
RNAcode5	1435086-1435148 (+) 1435113-1435148 (+)	0/0	-
RNAcode6 (*sX6*)	**1971528-1971770 (+)** 1971591-1971770 (+)	**15/1**	**1971505**
RNAcode7	3457704-3457808 (+) 3457734-3457808 (+)	0/0	-
RNAcode8	3463701-3463796 (+) 3463743-3463796 (+)	3/4	-
RNAcode9	3727337-3727555 (+) 3727361-3727555 (+) 3727445-3727555 (+)	21/1	3727405
RNAcode10	3848998-3849048 (+)	1/0	-
RNAcode11	4083126-4083464 (+) 4083348-4083464 (+)	0/0	-
RNAcode12	4368388-4368489 (+)	0/4	-
RNAcode13	4545859-4545918 (+)	0/0	-
RNAcode14	5098317-5098562 (+) 5098473-5098562 (+)	0/0	-
RNAcode15	644457-644513 (-)	0/0	-
RNAcode16	1219379-1219462 (-) 1219379-1219450 (-)	0/0	-
RNAcode17	1360037-1360129 (-) 1360037-1360120 (-)	0/0	-
RNAcode18	2429470-2429610 (-) 2429470-2429568 (-)	0/0	-
RNAcode19	2602554-2602655 (-)	0/0	-
RNAcode20	3215182-3215274 (-) 3215182-3215223 (-)	1/0	-
RNAcode21	3408680-3408883 (-)	4/0	3408906
RNAcode22	3881142-3881246 (-)	2/0	-
RNAcode23	4104001-4104081 (-) 4104001-4104060 (-)	1/1	-
RNAcode24	5054413-5054448 (-)	1/0	-

[a], number of reads (library 2/library 1) mapped to the respective loci. [b], automatically annotated TSS (see Table S2); "-" no TSS identified.

References

1. Thieme, F., Koebnik, R., Bekel, T., Berger, C., Boch, J., Büttner, D., Caldana, C., Gaigalat, L., Goesmann, A., Kay, S. *et al.* (2005) Insights into genome plasticity and pathogenicity of the plant pathogenic bacterium *Xanthomonas campestris* pv. vesicatoria revealed by the complete genome sequence. *J. Bacteriol.*, **187**, 7254-7266.
2. Washietl, S., Findeiß, S., Müller, S.A., Kalkhof, S., von Bergen, M., Hofacker, I.L., Stadler, P.F. and Goldman, N. (2011) RNAcode: Robust discrimination of coding and noncoding regions in comparative sequence data. *RNA*, **17**, 578-594.

Table S8. Genome sequences used for the calculation of multiple sequence alignments.

NCBI ID	Species name
NC_008752.1	*Acidovorax avenae* subsp. *citrulli* AAC00-1, complete genome
NC_004547.2	*Erwinia carotovora* subsp. *atroseptica* SCRI1043, complete genome
NC_010943.1	*Stenotrophomonas maltophilia* K279a, complete genome
NC_011071.1	*Stenotrophomonas maltophilia* R551-3, complete genome
NC_002516.2	*Pseudomonas aeruginosa* PAO1, complete genome
NC_009656.1	*Pseudomonas aeruginosa* PA7, complete genome
NC_008463.1	*Pseudomonas aeruginosa* UCBPP-PA14, complete genome
NC_005773.3	*Pseudomonas syringae* pv. *phaseolicola* 1448A, complete genome
NC_007005.1	*Pseudomonas syringae* pv. *syringae* B728a, complete genome
NC_004578.1	*Pseudomonas syringae* pv. *tomato* DC3000, complete genome
NC_003295.1	*Ralstonia solanacearum* GMI1000, complete genome
NC_003919.1	*Xanthomonas axonopodis* pv. *citri* 306, complete genome
NC_003902.1	*Xanthomonas campestris* pv. *campestris* ATCC 33913, complete genome
NC_007086.1	*Xanthomonas campestris* pv. *campestris* 8004, complete genome
NC_007508.1	*Xanthomonas campestris* pv. *vesicatoria* 85-10, complete genome
NC_002488.3	*Xylella fastidiosa* 9a5c, complete genome
NC_004556.1	*Xylella fastidiosa* Temecula1, complete genome
NC_006834.1	*Xanthomonas oryzae* pv. *oryzae* KACC10331, complete genome
NC_007705.1	*Xanthomonas oryzae* pv. *oryzae* MAFF 311018, complete genome

Table S9. The extended two-by-two confusion matrix summarizes the predictive power of the automated TSS annotation approach. A subsample of the manually curated TSS map of *H. pylori* (1) was used as reference data set and evaluated with our method. The analyzed data-set contained 392 manually annotated TSSs. In total 566 genomic positions fulfilled the criteria (accumulation of at least three read starts) to be analyzed by our automated TSS annotation approach. According to Fawcett (2006), true positives (classified as TSS and also manually annotated) and false positives (classified as TSS but not manually annotated) as well as true negatives (neither classified as TSS nor manually annotated) and false negatives (not classified as TSS but manually annotated) were evaluated (2). Based on these values the calculated values for sensitivity, specificity as well as positive and negative predictive values are listed.

	Manually annotated TSSs (392)		
Automatic classification (566)	True Positives: 321	False Positives: 124	Positive predictive value: 72%
	False Negatives: 71	True Negatives: 50	
	Sensitivity: 82%	Specificity: 29%	

References

1. Sharma, C.M., Hoffmann, S., Darfeuille, F., Reignier, J., Findeiss, S., Sittka, A., Chabas, S., Reiche, K., Hackermüller, J., Reinhardt, R. *et al.* (2010) The primary transcriptome of the major human pathogen *Helicobacter pylori*. *Nature*, **464**, 250-255.
2. Fawcett, T. (2006) An introduction to ROC analysis. *Pattern recognition letters*, **27**, 861-874.

Erklärung

Hiermit erkläre ich, dass ich die vorliegende wissenschaftliche Arbeit selbstständig und ohne fremde Hilfe verfasst habe. Ich erkläre weiterhin, dass andere als die von mir angegebenen Quellen und Hilfsmittel nicht benutzt und die den benutzten Werken wörtlich oder inhaltlich entnommenen Stellen als solche kenntlich gemacht wurden. Mit dieser Arbeit bewerbe ich mich erstmals um die Erlangung des Doktorgrades.

Weißenfels, 17. Dezember 2013

Cornelius Schmidtke

Lebenslauf

Persönliche Daten:

Name: Cornelius Schmidtke (geb. Schubert)

Anschrift: Lutherplatz 2 in 06667 Weißenfels

Geburtsdatum: 15.05.1982

Geburtsort: Jena

Familienstand: verheiratet

Staatsangehörigkeit: deutsch

Ausbildung:

1992-2000	Johann-Heinrich-Pestalozzi-Gymnasium Stadtroda
	Abschluss: Abitur
2000-2001	Zivildienst in Stadtroda
2001-2007	Studium der Biologie an der Martin-Luther-Universität Halle-Wittenberg
2006-2007	Diplomarbeit am Institut für Genetik, Abteilung Pflanzengenetik der Martin-Luther-Universität Halle-Wittenberg

        Thema: Analyse der potentiellen Virulenzfunktionen möglicher Adhäsine und Sekretionssysteme von *Xanthomonas campestris* pv. *vesicatoria*

        Abschluss: Diplom-Biologe (Note 1,1)

2007-2013   Promotionsarbeit am Institut für Biologie, Institutsbereich Genetik, Abteilung Pflanzengenetik der Naturwissenschaftlichen Fakultät I der Martin-Luther-Universität Halle-Wittenberg