Molecular Plant Pathology

MOLECULAR PLANT PATHOLOGY (2010) 11(6), 733–746 DOI: 10.1111/J.1364-3703.2010.00641.X

Microarray-based comparison of genetic differences between strains of *Streptomyces turgidiscabies* with focus on the pathogenicity island

MARJA AITTAMAA[1], PANU SOMERVUO[1], INTO LAAKSO[2], PETRI AUVINEN[3] AND JARI P. T. VALKONEN[1,*]

[1]*Department of Agricultural Sciences, PO Box 27, FIN-00014 University of Helsinki, Finland*
[2]*Division of Pharmaceutical Biology, Faculty of Pharmacy, PO Box 56, FIN-00014 University of Helsinki, Finland*
[3]*Institute of Biotechnology, PO Box 56, FIN-00014 University of Helsinki, Finland*

SUMMARY

The areas of the pathogenicity island (PAI) designated as 'colonization region' (CR) and 'toxicogenic region' (TR) [Lerat *et al*. (2009) *Mol. Plant Pathol*. **10**, 579–585] contain genes required for virulence and phytoxin production, respectively, in *Streptomyces* spp. causing common scab on potatoes. The PAI was tested for genetic variability by microarray analysis in strains of *S. turgidiscabies* isolated from potatoes in Finland. The data revealed four types of PAI based on divergent CR and TR which occurred in different combinations. Only one PAI type was highly similar to *S. scabies* (strains 87.22 and ATTC49173). Using probes designed for the predicted genes of *S. scabies*, two gene clusters in *S. scabies* appeared to be similar to most strains of *S. turgidiscabies* and contained PAI genes corresponding to CR and TR. They were located approximately 5 Mb apart in the *S. scabies* genome, as compared with only 0.3 Mb in *S. turgidiscabies* Car8. Data from comparative genomic hybridization with probes designed for *S. scabies* genes and for the PAI of *S. turgidiscabies* were compared by multilocus cluster analysis, which revealed two strains of *S. turgidiscabies* that were very closely related at the whole-genome level, but contained distinctly different PAIs. The type strain of *S. reticuliscabiei* (DSM41804; synonymous to *S. turgidiscabies*) was clustered with *S. turgidiscabies*. Taken together, the data indicate wide genetic variability of PAIs among strains of *S. turgidiscabies*, and demonstrate that PAI is made up of a mosaic of regions which may undergo independent evolution.

INTRODUCTION

Streptomyces species are Gram-positive bacteria which constitute a monophyletic clade, but possess wide genetic variability (Guo *et al*., 2008). A few are soil-borne plant pathogens, with those infecting potato being widely studied because of the significant yield losses caused (Loria *et al*., 2006; Naito *et al*., 2004; Wanner, 2009). *Streptomyces scabie*s was the first described pathogen causing common scab on potato tubers, and is distributed widely in potato-growing areas (Lambert and Loria, 1989a; Lerat *et al*., 2009). More recently, many additional *Streptomyces* species have been found to cause common scab on potatoes (Bouchek-Mechiche *et al*., 2000a; Goyer *et al*., 1996; Lambert and Loria, 1989b; Miyajima *et al*., 1998; Park *et al*., 2003; Wanner, 2009). One is *Streptomyces turgidiscabies*, isolated and characterized from potato tubers in Japan (Miyajima *et al*., 1998; Takeuchi *et al*., 1996) and Finland (Kreuze *et al*., 1999; Lindholm *et al*., 1997), and later found to be widely spread and important in potato crops in many other parts of the world (reviewed in Lehtonen *et al*., 2004; Naito *et al*., 2004; Wanner, 2009). In Finland, *S. scabies* and *S. turgidiscabies* are found in the same fields, potato tubers and even in the same scab lesions (Lehtonen *et al*., 2004). Pathogenic strains of *S. turgidiscabies* tolerate lower pH than do *S. scabies*, which is beneficial for the bacterium in areas such as Finland where precipitation exceeds evaporation, resulting in leaching and acidification of soils (Lindholm *et al*., 1997). Furthermore, pathogenic strains of *S. turgidiscabies* can be antagonistic to *S. scabies* (Hiltunen *et al*., 2009). These data suggest that *S. turgidiscabies* competes for the ecological niche with *S. scabies* and has the potential to become more common as the cause of common scab in northern Europe.

The symptoms of common scab are caused by thaxtomin A, a phytotoxin synthesized by both *S. turgidiscabies* and *S. scabies* (Goyer *et al*., 1998; Hiltunen *et al*., 2006; Lawrence *et al*., 1990; Loria *et al*., 2008; Toth *et al*., 1998). The synthesis of thaxtomin A is induced by cell wall polymer subunits, such as cellobiose and cellotriose (Johnson *et al*., 2007; Joshi *et al*., 2007a), and is thought to interfere with cell wall synthesis in plants (Scheible *et al*., 2003). Thaxtomin A is a cyclic dipeptide consisting of a nitrotryptophan and a phenylalanine linked by the action of a

**Correspondence*: Email: jari.valkonen@helsinki.fi

© 2010 THE AUTHORS
MOLECULAR PLANT PATHOLOGY © 2010 BSPP AND BLACKWELL PUBLISHING LTD

nonribosomal peptide synthase operon that includes genes *txtA* and *txtB* (Healy *et al*., 2000). In the dipeptide backbone, two hydroxyl groups are added to phenylalanine by a P450 monooxygenase encoded by *txtC* (Healy *et al*., 2002). Close to this operon resides the gene *txtR*, a positive transcriptional regulator of the thaxtomin synthesis genes (Joshi *et al*., 2007a; Kers *et al*., 2005). Thaxtomin A is nitrated at the 4-position probably by a P450 monooxygenase encoded by a gene located next to a nitric oxide synthase gene (*nos*) (Johnson *et al*., 2008). These two genes are linked to *txtR* and the txtABC operon, but separated from them by a few putative genes of unknown function (Kers *et al*., 2005). They belong to the pathogenicity island (PAI) of *S. turgidiscabies* characterized in strain Car8 isolated from carrot in Japan (Bukhalid *et al*., 1998; Kers *et al*., 2005). This strain causes common scab symptoms also on potato tubers (Bouchek-Mechiche *et al*., 2006).

Three PAI regions of strain Car8 have been sequenced (Kers *et al*., 2005) and are further referred to as PAI-I, PAI-II and PAI-III. The thaxtomin synthesis genes discussed above are located in PAI-III which is homologous also in *S. scabies*, *S. acidiscabies* (Kers *et al*., 2005) and probably most, if not all, *Streptomyces* species that cause common scab (Loria *et al*., 2006), as suggested by the detection of *txtA* and *txtB* and thaxtomin production (Bouchek-Mechiche *et al*., 2006; Wanner, 2009). Next to this segment of PAI-III, *S. turgidiscabies* contains a plant fasciation (fas) operon constituted by six *fas* genes and two methyltransferase genes, flanked by IS110 family transposable elements and able to induce gall formation in *Arabidopsis thaliana* and *Nicotiana tabacum* (Joshi and Loria, 2007). The fas operon is not detected in *S. scabies* and *S. acidiscabies* (Kers *et al*., 2005).

The other two regions, PAI-I and PAI-II, are next to each other at a distance of *c.* 0.3 Mb from PAI-III, and were localized in strain Car8 using the gene *nec1* as a probe (Kers *et al*., 2005). *nec1* is situated in PAI-II and encodes a secreted protein that contributes to virulence, but is not necessary for pathogenicity in *S. turgidiscabies* (Joshi *et al*., 2007b). PAI-I contains the gene *tomA*, which has been studied in *S. scabies* (strain 87.22). It encodes tomatinase, which can detoxify the glycoalkaloid α-tomatine, but has little significance to the virulence of *S. scabies* in tomato plants (Seipke and Loria, 2008). Although *nec1* and *tomA* are detected in a wide range of common scab-inducing *Streptomyces* species, dispensability for the pathogenicity of these genes (Joshi *et al*., 2007b; Seipke and Loria, 2008) is evident from the existence of pathogenic strains which lack one or both genes (Kreuze *et al*., 1999; Wanner, 2009).

The low G + C content of *nec1* (54%) (Kers *et al*., 2005), when compared with the high overall estimated G + C content of the *S. turgidiscabies* genome (71%) (Miyajima *et al*., 1998), suggests that the origin of *nec1* is different from that of the genus *Streptomyces* (Joshi *et al*., 2007b), which is also supported by its linkage to a transposon pseudogene (Bukhalid *et al*., 1998). In addition, many other genes among the 31 putative genes detected in the sequenced regions of the PAI have a G + C content lower than 71%. It has been suggested that *nec1* defines a mobile element that can move horizontally between plant-pathogenic *Streptomyces* species (Bukhalid *et al*., 2002; Healy *et al*., 1999). Indeed, the genomic segment containing the characterized PAI regions I–III can move from *S. turgidiscabies* to *S. diastatochromogenes* and *S. coelicolor* (and further from *S. coelicolor* to *S. lividans*) during conjugal mating, which allows *S. diastatochromogenes*, but not *S. coelicolor*, to acquire pathogenic properties on plants (Kers *et al*., 2005). Hence, the genetic background of the recipient species or strain influences the expression of pathogenicity and virulence (Kers *et al*., 2005). The other *Streptomyces* species causing common scab (Bouchek-Mechiche *et al*., 2000a; Goyer *et al*., 1996; Lambert and Loria, 1989b; Miyajima *et al*., 1998; Park *et al*., 2003) were detected and described much later than *S. scabies* (Lambert and Loria, 1989a; Lerat *et al*., 2009), which has resulted in a hypothesis suggesting that the latter described species would have gained the PAI from *S. scabies* (Healy *et al*., 1999; Loria *et al*., 2006). It is one possibility, but is supported by relatively limited experimental evidence. With regard to studies on *S. turgidiscabies*, previous molecular studies on the genetic variability have focused on just a few genes of the PAI (Bukhalid *et al*., 1998, 2002; Healy *et al*., 1999), and have used a physiologically and phenotypically homogeneous population of strains isolated in Hokkaido Island, Japan, possibly originating from a single introduction (Miyajima *et al*., 1998; Takeuchi *et al*., 1996). Therefore, it is important to extend the molecular genetic analysis of *S. turgidiscabies* to include strains occurring in areas such as northern Europe, where *S. turgidiscabies* shows considerable variability in terms of virulence and the presence or absence of *nec1* (Hiltunen *et al*., 2005; Kreuze *et al*., 1999; Lehtonen *et al*., 2004; Lindholm *et al*., 1997).

The sequence data available on bacterial genomes can be efficiently utilized in microarray-based analyses. The genetic variability of 15 *Streptomyces* species has been studied using microarrays with probes consisting of PCR fragments amplified from over 7500 open reading frames (ORFs) of *S. coelicolor* (Hsiao and Kirby, 2008, 2009). However, plant-pathogenic *Streptomyces* species were not included, partly because the genome sequence of *S. scabies* strain 87.22 has become available only recently (GENBANK accession no. FN554889). In a previous study (Aittamaa *et al*., 2008), the draft genome sequence of *S. scabies* was used to design specific probes for 3919 ORFs of the more than 8000 putative ORFs of *S. scabies*, and also 113 probes for the 31 ORFs and their intergenic regions in the PAI of *S. turgidiscabies* Car8. The custom-designed probes for the *Streptomyces* spp. and other bacteria were synthesized on the array and used to distinguish bacterial pathogens of potato (Aittamaa *et al*., 2008). The aim of the present study was to carry out a more detailed comparison of the genetic differences between

© 2010 THE AUTHORS
 MOLECULAR PLANT PATHOLOGY © 2010 BSPP AND BLACKWELL PUBLISHING LTD

Table 1 *Streptomyces* strains used in this study.

Strain*	Species	16S rRNA and IGS sequences†	Place of isolation	Potato cultivar	Pathogenicity on potato‡	Thaxtomin A production§
32	*S. turgidiscabies*	EU828541	Norra Sunderbyn, Sweden	Bintje	Yes[1,2,3]	Yes
255	*S. turgidiscabies*	EU828535	Apukka, northern Finland	Sabina	No[6]	No
261	*S. turgidiscabies*	EU828538	Apukka, northern Finland	Amazone	Yes[1,6]	Yes
287	*S. turgidiscabies*	Y15495, EU828529	Mikkeli, south-east Finland	Fambo	Yes[1,4,6]	Yes
300	*S. turgidiscabies*	Y15499, EU828531	Mikkeli, south-east Finland	Rocket	Yes[1,2,4,6]	Yes
304	*S. turgidiscabies*	EU828539	Partala, south-east Finland	Matilda	No[1,6]	No
323	*S. turgidiscabies*	EU828540	Tyrnävä, northern Finland	Timo	Yes[1,6]	Yes
342	*S. turgidiscabies*	EU828536	Tyrnävä, northern Finland	Fambo	Yes[1,6]	Yes
368	*S. turgidiscabies*	EU828537	Sotkamo, north-east Finland	Matilda	Yes[1,6]	Yes
DSM 41804	*S. reticuliscabiei,* type strain	AJ007428, AY296981	France	Unknown	Yes[5]	No
ATCC 49173	*S. scabies,* type strain	AB026199	Maine, USA	Unknown	Yes[1]	Yes
250	*Streptomyces* sp.	EU828546	Apukka, Finland	Van Gogh	No[6]	No
266	*Streptomyces* sp.	EU828542	Apukka, Finland	Satu	No[6]	No
286	*Streptomyces* sp.	EU828547	Mikkeli, Finland	Fambo	No[1,6]	No
300B	*Streptomyces* sp.	EU828545	Unknown	Unknown	No[1]	No
Additional reference strains:						
14	*S. scabies*	EU828544	Rasmyran, Sweden	Bintje	Yes[2,3]	Yes
208	*S. scabies*	EU828543	Robersfors, Sweden	Eloge	Yes[3]	Yes
267	*S. scabies*	Y15500, EU828533	Sotkamo, Finland	Van Gogh	Yes[6]	Yes
289	*S. scabies*	Y15509, EU828534	Mikkeli, Finland	Felsina	Yes[6]	Yes
364	*S. scabies*	Y15497, EU828532	Tyrnävä, Finland	Fambo	Yes[2,4,6]	Yes

*For species' identification, see Kreuze *et al*. (1999), Lehtonen *et al*. (2004) and Fig. 3. *Streptomyces reticuliscabiei* was obtained from the German Collection of Microorganisms and Cell Cultures (DSM). Also, strains 287, 300 and 364 are available from DSM under the accession numbers DSM41745, DSM41747 and DSM41744, respectively. ATCC, American Type Culture Collection. The additional reference strains of *S. scabies* described previously (Aittamaa *et al*., 2008) were included in phylogenetic comparisons only.

†The National Center for Biotechnology Information (NCBI) sequence database accession numbers for 16S rRNA gene and internal genomic region (IGS) sequences, respectively; alternatively, both sequences included in a single database accession.

‡Pathogenicity was tested by growing healthy potato plants in disinfected and subsequently inoculated soil in this study (1), by Hiltunen *et al*. (2005) (2), Lehtonen *et al*. (2004) (3), Kreuze *et al*. (1999) (4) and Bouchek-Mechiche *et al*. (2000b) (5), or was tested by inoculation of immature potato minitubers (Lindholm *et al*., 1997) (6).

§Thaxtomin A detected by thin layer chromatography and high-pressure liquid chromatography. Strains 304 and 255 produced yellow substances that were not thaxtomin A (see Fig. S1).

strains of *S. turgidiscabies* using the probes designed for *S. turgidiscabies* and *S. scabies*, and by including the type strains of *S. scabies* and *S. reticuliscabiei* for comparison. *Streptomyces reticuliscabiei* was included because it is a potato pathogen taxonomically synonymous to *S. turgidiscabies* (Bouchek-Mechiche *et al*., 2006), but was originally named differently owing to the symptoms of netted scab rather than common scab which it induces on potato tubers (Bouchek-Mechiche *et al*., 2000a, b). The main focus of our study was on the genetic variability of the three PAI regions whose sequences are available from *S. turgidiscabies*.

RESULTS

The eight strains of *S. turgidiscabies* from Finland and one strain of *S. turgidiscabies* from northern Sweden included in the study were originally isolated from common scab lesions of potatoes grown in the field. These strains have been tested for pathogenicity in previous studies, and most were tested again in this study to reconfirm their virulence or avirulence (Table 1). Results showed that seven strains caused common scab lesions on potato tubers. These strains also produced thaxtomin A *in vitro* (Table 1). In contrast, two strains (255 and 304) did not induce symptoms on potato tubers (Table 1) and also did not produce thaxtomin A, but other substances were detected by thin layer chromatography (TLC) and high-pressure liquid chromatography (HPLC) (Fig. S1, see Supporting Information). The type strain of *S. scabies* (ATCC49173) induced scab symptoms and produced thaxtomin, whereas no thaxtomin production was detected with the type strain of *S. reticuliscabiei* (DMS41804) that induces netted scab symptoms distinct from common scab (Bouchek-Mechiche *et al*., 2000b). Moreover, the four nonpathogenic strains of unknown *Streptomyces* species included for comparison produced no detectable amounts of thaxtomin A (Table 1).

Genetic differences of *S. turgidiscabies* strains detected with probes designed for the PAI genes of *S. turgidiscabies* strain Car8

The microarray included one to five probes for each of the 31 ORFs detected in the three characterized PAI regions of *S. turgi-*

© 2010 THE AUTHORS

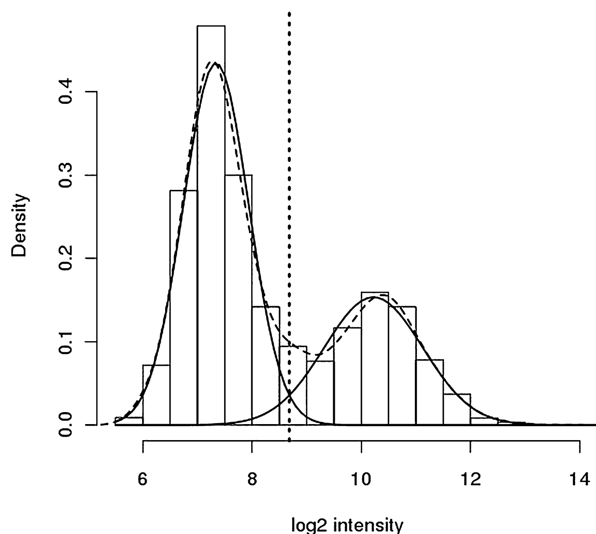


Fig. 1 A Gaussian mixture model applied to hybridization signal intensity values generated by microarray analysis of *Streptomyces turgidiscabies*, strain 287. Logarithmic intensities of 4032 probes, including 3919 and 113 probes designed on the basis of the open reading frames (ORFs) of *S. scabies* and the pathogenicity island (PAI) region of *S. turgidiscabies*, respectively, were used to estimate the parameters. The bar plot shows the histogram and the broken line indicates the distribution based on kernel density analysis. The full line depicts the two-component Gaussian mixture model decomposition. The vertical broken line shows the Bayesian classification boundary which was analytically computed on the basis of the estimated parameters of the mixture model. The two signal categories 0 and 1 correspond to the left (weak signals) and right (strong signals) sides of the classification boundary. The model parameters are hybridization-specific. Signal categorization was performed separately for each strain. The overlapping area of the two Gaussians is 0.03, indicating that only 3% of the probes are incorrectly classified.

discabies strain Car8, except for the short ORF25 (*stPAI025*), for which no probe with sufficient target specificity could be obtained. All available sequence data of strain Car8 (Kers *et al*., 2005) were taken into account in probe design. Signal intensities were categorized into two classes based on an array-specific Gaussian mixture model (Asyali *et al*., 2004) (Fig. 1). Accordingly, each signal was categorized either to category 0 (weak signal) or category 1 (strong signal). After repetition of the hybridization experiment, the signal category for each probe and sample was averaged. The averaged signals 0 and 1 represent reproducible categories. In contrast, if the signal of a probe was classified to category 0 in one experiment and to category 1 with the same sample in another experiment, the resulting signal category for this probe and sample would be 0.5. Signal intensities presented using these categories underline the major differences between strains (Fig. 1). Depending on the hybridization and strain of *S. turgidiscabies*, the overlapping area of the two Gaussians was 0.01–0.06, indicating that only 1%–6% of the probes were classified incorrectly.

The sequenced regions PAI-I and PAI-II (Fig. 2) are located within *c*. 10 kb in the PAI of *S. turgidiscabies* Car8 (Kers *et al*., 2005). The analysis of five of the seven pathogenic strains of *S. turgidiscabies* (32, 261, 323, 342 and 368) resulted in category 1 signals from all probes for the three and six ORFs of PAI-I and PAI-II, respectively (Fig. 2b). These probes gave category 1 signals also with *S. scabies* ATCC49173 and the two strains of *S. turgidiscabies* (304 and 255) that caused no scab symptoms on potatoes and did not produce detectable amounts of thaxtomin A. In contrast, signals for PAI-I and PAI-II were mostly classified to category 0 with the pathogenic strains 287 and 300 of *S. turgidiscabies*. The pathogenic strain DMS41804 of *S. reticuliscabiei* gave mostly category 0 signals for PAI-II and also with many probes designed for PAI-I of Car8. The four nonpathogenic strains of unknown *Streptomyces* species (300B, 250, 268 and 286) were mostly characterized by category 0 signals with the probes of PAI-I and PAI-II (Fig. 2b).

Hybridization with three pathogenic strains of *S. turgidiscabies* (32, 287 and 323) generated category 1 signals with all of the probes for PAI-III. In addition, other pathogenic strains of *S. turgidiscabies* gave mostly category 1 signals with probes for the txtABC gene cluster and also *nos*, whereas hybridization with *S. reticuliscabiei* and the four unknown *Streptomyces* species resulted mainly in category 0 signals (Fig. 2b).

The data obtained by polymerase chain reaction (PCR) tests with primers designed to the 31 PAI genes of *S. turgidiscabies* (strain Car8) (Tables S1 and S2, see Supporting Information) were largely consistent with the binary signal categorization of microarray hybridization intensities (Fig. 2). The few discrepancies were not unexpected because the microarray probes and PCR primers were designed to different positions of ORFs as a result of the criteria stipulated in the respective probe/primer design programmes. Hence, discrepancies could be a result of sequence divergence, which might be higher for some parts of the ORF than others between the strains tested and the strain Car8 used for probe and primer design. However, the overall signal patterns and PCR results were very similar, which supports the microarray signal categorization used in this study.

The thaxtomin synthesis genes (txt) showed sequence variability, as revealed by microarray analysis. Therefore, the regions of txt genes targeted by some probes were amplified and sequenced from strains 32, 287, 300 and 304 of *S. tugidiscabies* and *S. scabies* ATCC49173. The level of heterogeneity between the 40-nucleotide-long, target-specific part of the 60-nucleotide-long probes and the corresponding target sequences in the txt genes was exemplified using the region of the *txtA* probe which gave category 0 signals with some strains and category 1 signals with others (Fig. 2c). The target sequence in strains 32 and 287 contained no mismatch with the probe, and these strains gave category 1 signals (Fig. 2c). In contrast, the target sequences of strains 300 and 304 contained seven mismatches (identity

© 2010 THE AUTHORS
 MOLECULAR PLANT PATHOLOGY © 2010 BSPP AND BLACKWELL PUBLISHING LTD

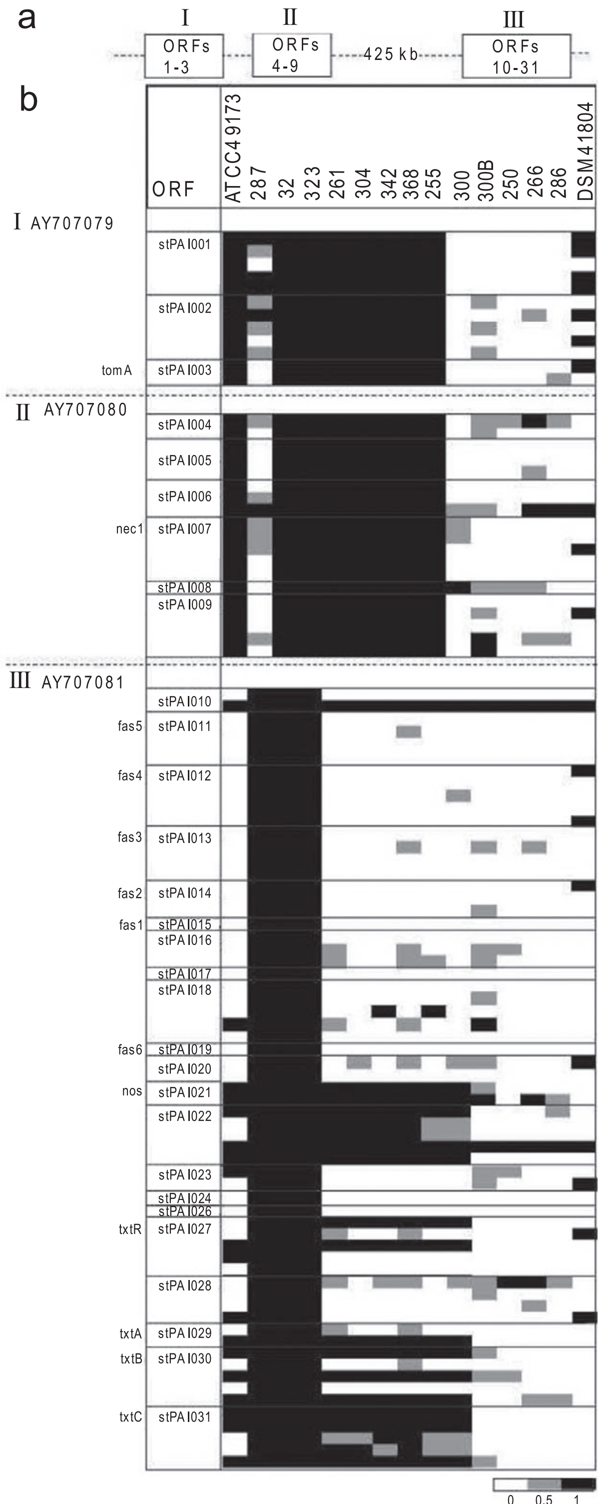
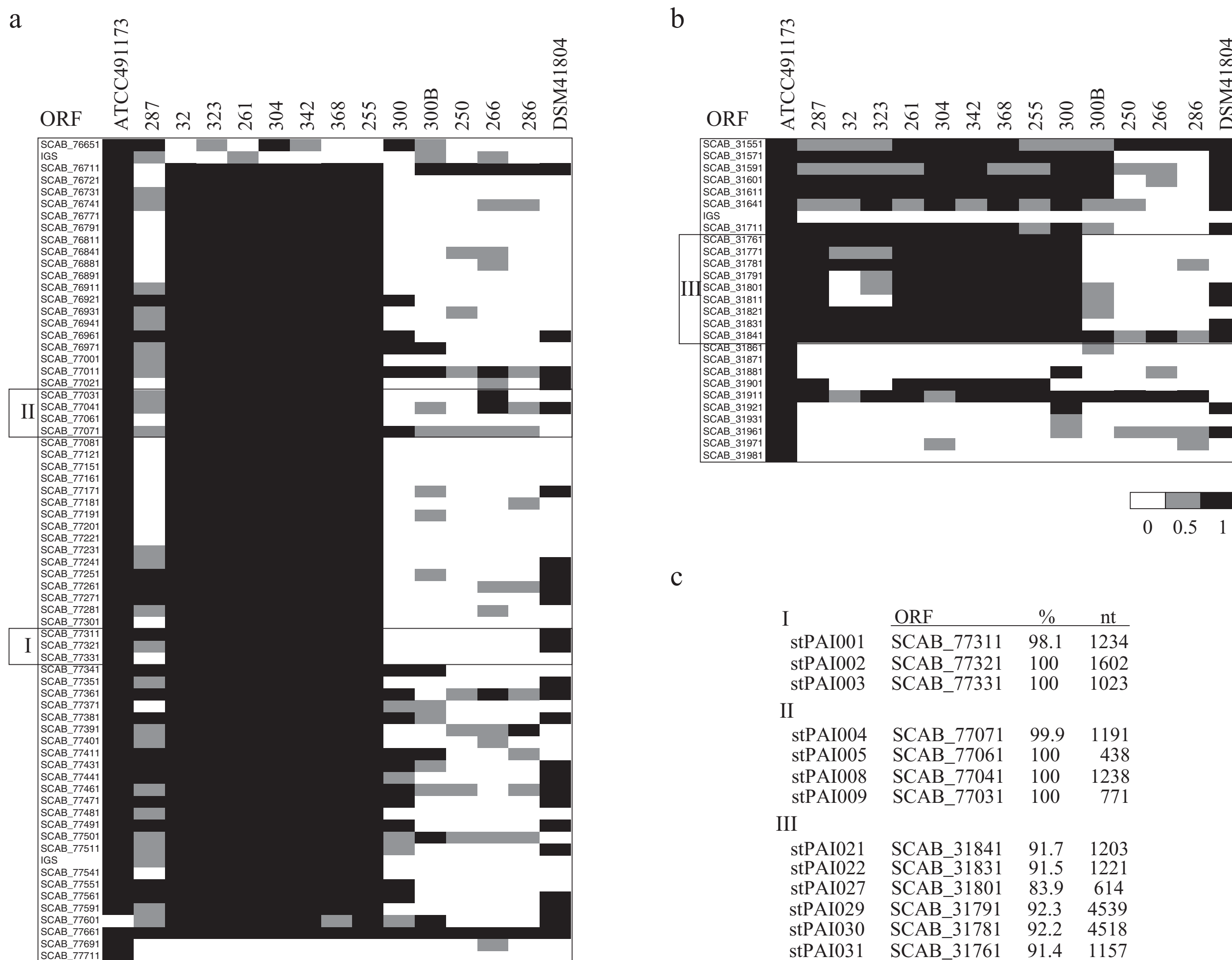


Fig. 2 Microarray analysis of genomic DNA of *Streptomyces* strains using probes designed to the open reading frames (ORFs) in the pathogenicity island (PAI) of *Streptomyces turgidiscabies*. Seven pathogenic strains (287, 32, 323, 261, 342, 368 and 300) and two nonpathogenic strains (304 and 255) of *S. turgidiscabies*, three strains (250, 266 and 286) of unknown *Streptomyces* species isolated from potato scab lesions, one *Streptomyces* strain (300B) of unknown origin, and the pathogenic type strains of *S. scabies* (ATCC49173) and *S. reticuliscabiei* (DSM41804) were tested, and the combined data from two experiments are shown. (a) Schematic presentation of the three sequenced PAI regions (I, II and III) of *S. turgidiscabies* (accession numbers AY707079, AY707080 and AY707081, respectively) (Kers *et al*., 2005). (b) Hybridization intensities of one to five probes per ORF using intensity categorization into two classes based on a two-component Gaussian mixture model (see Fig. 1) and averaging the results of two independent hybridizations, which resulted in three possible values: 0 (weak signal), 0.5 or 1 (strong signal). The names of the ORFs (genes) characterized in previous studies are indicated (see text for details). (c) Sequence alignment illustrating nucleotide differences along one of the probes designed for txtA [topmost probe of stPAI029 in (b)] of *S. turgidiscabies* strain Car8 (Kers *et al*., 2005). The sequences of strains 32 and 287 of *S. turgidiscabies* are identical to the probe, consistent with category 1 hybridization signals. In contrast, strains 300 and 304 of *S. turgidiscabies* and strain ATCC49173 of *S. scabies* (Ss) contain mismatches, which caused a loss of hybridization intensity and a category 0 signal.

© 2010 THE AUTHORS



	ORF	%	nt
I			
stPAI001	SCAB_77311	98.1	1234
stPAI002	SCAB_77321	100	1602
stPAI003	SCAB_77331	100	1023
II			
stPAI004	SCAB_77071	99.9	1191
stPAI005	SCAB_77061	100	438
stPAI008	SCAB_77041	100	1238
stPAI009	SCAB_77031	100	771
III			
stPAI021	SCAB_31841	91.7	1203
stPAI022	SCAB_31831	91.5	1221
stPAI027	SCAB_31801	83.9	614
stPAI029	SCAB_31791	92.3	4539
stPAI030	SCAB_31781	92.2	4518
stPAI031	SCAB_31761	91.4	1157

Fig. 3 Microarray analysis of genomic DNA of *Streptomyces* strains using probes designed to *Streptomyces scabies* (strain 87.22) whose genome sequence has been determined (FN554889). Most probes were designed to putative open reading frames (ORFs), but some are for intergenic regions (IGS). The probes in two genomic regions of *S. scabies* (nucleotides 8 477 342–8 587 834 and 3 598 108–3 616 372) gave category 1 signals for most strains of *S. turgidiscabies* tested (averaged signal categories from replicated hybridizations). These genomic regions of *S. scabies* contain clusters of ORFs (boxed) nearly identical to the ORFs in the pathogenicity islands (PAI) I, II and III in *S. turgidiscabies* (strain Car 8), as indicated in (c) by comparison of nucleotide sequence identities (%). nt, the length of the gene sequence (nucleotides) used for pairwise comparison. ATCC49173 and DSM41804 are type strains of *S. scabies* and *S. reticuliscabiei*, respectively. Strains 300B, 250, 266 and 286 are unknown *Streptomyces* species, also isolated from potato scab lesions as the other strains tested.

82.5%) and *S. scabies* ATCC49173 four mismatches (identity 90%) as compared with the probe; these strains gave category 0 signals (Fig. 2c). This validation also supported the microarray signal categorization used in this study.

Comparison of *S. turgidiscabies* strains and *Streptomyces* species using probes designed for genes of *S. scabies* strain 87.22

Probes for 3919 ORFs of *S. scabies* strain 87.22 spanning the whole genome (FN554889) were used to test the strains of *S. turgidiscabies* and the type strains of *S. scabies* and *S. reticulicabiei*. Category 1 signals were generated by probes for *S. scabies* genes forming two clusters situated *c*. 5 Mb apart in the genome of strain 87.22. These clusters contained ORFs similar to the *S. turgidiscabies* PAI (stPAI) genes (Fig. 3). The larger gene cluster covered a region of *c*. 100 kb and could be defined by nucleotides 8 477 342–8 587 834 (putative ORFs SCAB_76 711–77 661) in the available genome sequence of *S. scabies*, because the probes for all genes of this region gave category 1 signals for seven strains of *S. turgidiscabies* (Fig. 3a). Probes for genes in regions flanking this genomic segment of *S. scabies* showed category 0 signals for strains of *S. turgidiscabies*. This gene cluster contained genes similar to *stPAI001–stPAI003* (PAI-I) and *stPAI004*, *stPAI005*, *stPAI008* and *stPAI009* (PAI-II).

© 2010 THE AUTHORS

The smaller gene cluster whose probes gave category 1 signals for *S. turgidiscabies* covered a region of *c.* 10 kb in *S. scabies* and contained genes similar to stPAI-III. It was defined by nucleotides 3 598 108–3 616 372 (ORFs SCAB_31 761–31 841) in the genome of strain 87.22 based on criteria similar to those explained above (Fig. 3b). This genomic region of *S. scabies* contains genes similar to the txt genes (*stPAI027*, *stPAI029*, *stPAI030* and *stPAI031*), *nos* (*stPAI021*) and the putative P450 cytochrome gene (*stPAI022*) of *S. turgidiscabies* Car8 (Fig. 3b).

Streptomyces turgidiscabies strains 32, 323, 261, 304, 342, 368 and 255 gave category 1 signals with probes for the *S. scabies* ORFs constituting the gene cluster which contains genes similar to stPAI-I + II; however, only a few of these probes gave a category 1 signal for strain 287 (Fig. 3a). On the other hand, the probes for *S. scabies* ORFs which formed a gene cluster similar to stPAI-III generated category 1 signals for other strains of *S. turgidiscabies*, except strains 32 and 323. Strain 300 of *S. turgidiscabies* was exceptional in that very few probes for either region gave category 1 signals, which was also the case with the unknown nonpathogenic *Streptomyces* strains. Roughly one-half of these probes gave category 1 signals for *S. reticulisabiei*. However, nearly all probes gave category 1 signals for *S. scabies* type strain ATCC49173, which suggests that this strain and strain 87.22 of *S. scabies* are quite similar for these genomic regions.

A total of 120 probes designed for ORFs of *S. scabies* gave category 1 signals with all strains of *S. scabies* and *S. turgidiscabies* tested and, furthermore, were >90% identical to ORF sequences of *S. coelicolor*, *S. avermitilis* and *S. griseus*, whose genome sequences are available (Hsiao and Kirby, 2008, 2009). Annotation of the corresponding ORFs in *S. coelicolor*, *S. avermitilis* and *S. griseus* showed that they were concordant with each other (Table S3, see Supporting Information). These probes may serve as a set of positive controls in future microarray studies on *Streptomyces*.

Clustering of S*treptomyces* strains on the basis of microarray and sequence data

The data from microarray analysis were used for clustering bacterial strains based on the probed genome sites that differed between strains as revealed by differences in signal intensities. The *p* distances were determined and the trees were drawn in MEGA4 (Kumar *et al*., 2008) using signals from the 113 probes designed for the 31 PAI genes of *S. turgidiscabies* (Fig. 4a) and the probes for 3919 putative ORFS of *S. scabies* (Fig. 4b). Comparison of the two trees revealed that the relationships inferred from PAI gene analysis (Fig. 4a) were quite consistent with the overall relatedness of the whole genomes of strains (Fig. 4b). However, in the clustering based on the PAI probes (Fig. 4a), *S. scabies* ATCC49173 was located close to the branch consisting of five *S. turgidiscabies* strains (Fig. 4a).

Strains 287 and 304 showed a striking contradiction in their clustering based on the PAI genes and in terms of the genomic similarities based on analysis with *S. scabies* probes. In the clustering based solely on the PAI probes, they were in separate clusters (Fig. 4a), whereas analysis with the 3919 *S. scabies* probes placed these two strains close to each other (Fig. 4b). These data provide compelling evidence for independent origins or different evolution of the PAI regions in strains 287 and 304 of *S. turgidiscabies*, despite their close relatedness and origin in the same area of south-east Finland.

The 16S rRNA gene sequences (Table 1) were subjected to neighbour-joining analysis to elucidate the phylogenetic relationships of the bacterial strains. A few additional strains of *S. scabies* used in our previous study (Aittamaa *et al*., 2008) were included for comparison (Table 1). Grouping of strains in the phylogenetic tree was mainly consistent with their previously determined taxonomic status (Fig. 4c). Strains of *S. turgidiscabies* and *S. scabies* were each placed in a distinct cluster. The type strain of *S. reticuliscabiei* was placed in the main cluster of *S. turgidiscabies*, as expected (Bouchek-Mechiche *et al*., 2006). Phylogenetic clustering based on the intergenic regions of ribosomal rRNA genes (Fig. 4d) was consistent with the tree based on 16S rRNA gene sequences, but revealed a wider range of genetic variability among the strains. For example, strains 287, 300 and 304 originating in the same geographical location in south-east Finland (Lindholm *et al*., 1997) were placed in a subcluster separated from other *S. turgidiscabies* strains. Furthermore, the data suggested only a distant relatedness between *S. scabies* ATCC49173 and the Finnish and Swedish strains of *S. scabies* (Fig. 4d).

DISCUSSION

Many *Streptomyces* species produce significant yield losses in potato production by causing common scab symptoms on tubers. They are of increasing concern because they are common in many countries in different parts of the world (e.g. Bouchek-Mechiche *et al*., 2000a; Goyer *et al*., 1996; Kreuze *et al*., 1999; Lambert and Loria, 1989a, b; Park *et al*., 2003; Takeuchi *et al*., 1996; Wanner, 2009). Their potential to obtain novel genetic material and to exchange pathogenicity and virulence factors via horizontal gene transfer seems obvious in the light of recent studies (Bukhalid *et al*., 2002; Kers *et al*., 2005). It is therefore important to develop tools that can capture molecular differences of common scab pathogens and their strains at multiple genetic loci simultaneously, ideally at the length of the whole genome.

In this study, the genetic variability of *S. turgisdiscabies* was analysed using a microarray-based approach, which allowed the efficient usage of sequence data available from the PAI of *S. turgisdiscabies* and also the whole genome sequence of *S. scabies*. The data revealed that the genetic variability in the

© 2010 THE AUTHORS

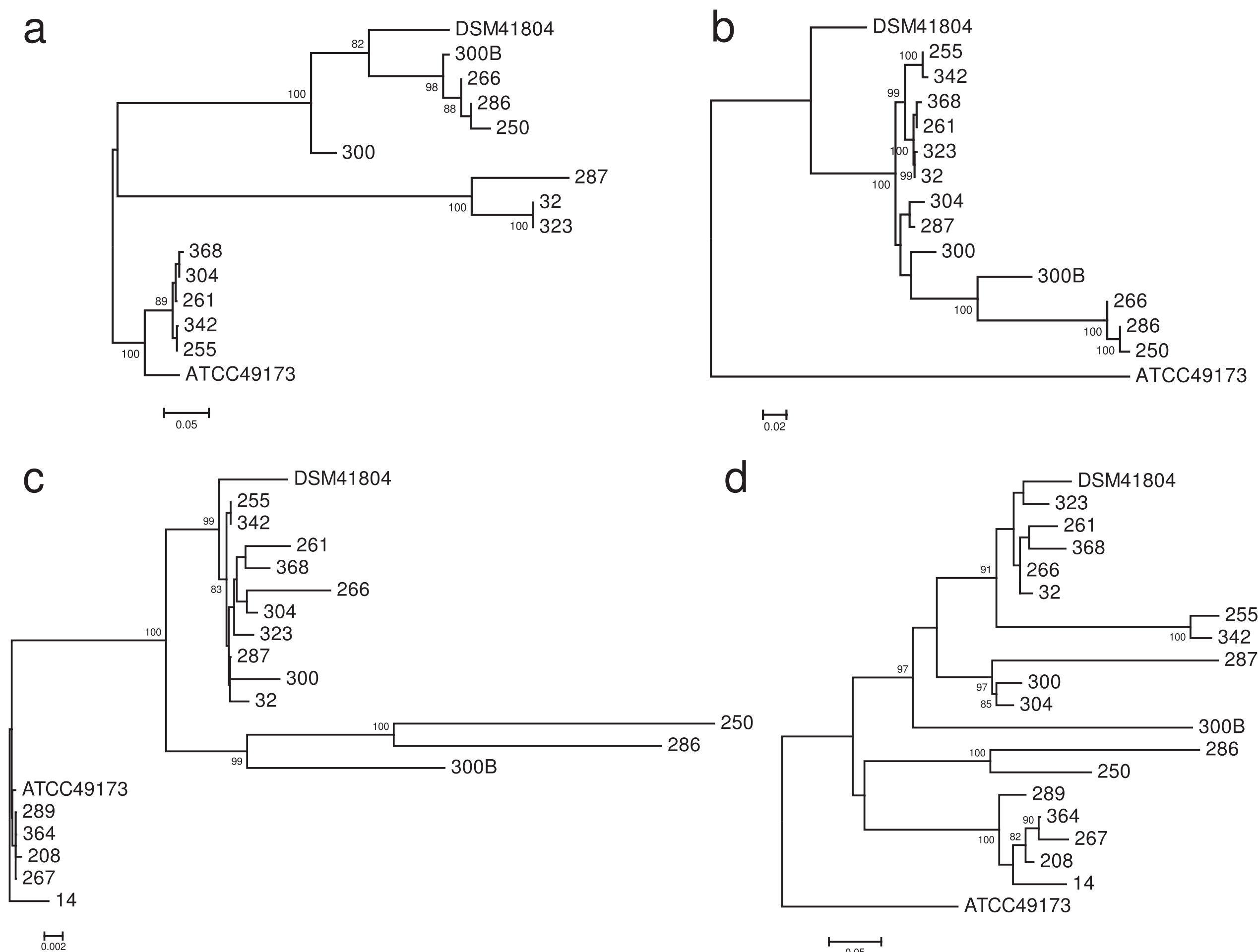


Fig. 4 Clustering of *Streptomyces* strains based on microarray analysis of genomic differences (*p* distances, a and b) and neighbour-joining analysis (c and d) of sequences of the ribosomal RNA cistron. Clustering based on signals detected with 113 probes designed on the basis of the 31 pathogenicity island (PAI) genes and their intergenic regions of *S. turgidiscabies* (strain Car8) (a) and probes for 3913 putative genes of *S. scabies* (strain 87.22) (b) (for signal categorization, see Fig. 1). Genetic distances determined on the basis of the 16S rRNA gene sequences (c) and the intergenic (IGS) sequences between the 16S and 23S rRNA genes (d). Trees (c) and (d) were constructed using the Kimura two-parameter model (Kimura, 1980). Bootstrap values were determined by 1000 replicates, and values greater than 80% are shown. Bars in (a) and (b) indicate the *p* distance, and those in (c) and (d) indicate Kimura nucleotide differences (K_{nuc}). Trees (c) and (d) include additional, previously characterized strains of *S. scabies* (14, 208, 267, 289 and 364) for comparison.

genomic regions containing pathogenicity-related genes is wide among strains of *S. turgidiscabies* in Finland, which contrasts with the previous conception of a highly homologous PAI in the strains of *S. turgidiscabies* from Hokkaido Island, Japan (Bukhalid *et al*., 1998; Miyajima *et al*., 1998). The previously sequenced regions in the PAI of the Japanese strain Car8 of *S. turgidiscabies* were designated as PAI-I + II and PAI-III in this study. They are located *c.* 0.3 Mb apart, encompass a total of *c.* 50 kb and contain 31 putative genes (Kers *et al*., 2005). Regions PAI-I + II contain nine genes and can be called a 'colonization region' (CR) (*sensu* Lerat *et al*., 2009) owing to genes such as *tom* and *nec1* which contribute to virulence, but are dispensable for pathogenicity (Joshi *et al*., 2007b; Seipke and Loria, 2008). Region PAI-III includes the genes needed for thaxtomin synthesis and can be called a 'toxicogenic region' (TR) (Lerat *et al*., 2009). The notations CR and TR are based on the analysis of genes in the PAI of *S. turgidiscabies* (strain Car8), but were also proposed for the corresponding regions of *S. scabies* (Lerat *et al*., 2009) in which the PAI (or PAIs), however, has not been characterized in closer detail. Indeed, because only a limited amount of sequence data are available from the PAI of *S. turgidiscabies* for comparison, the similarities of the PAIs in the two species are currently difficult to determine. The present study provides some novel information about the genes that might be homologous in the PAIs of the two species, which was possible to infer from data obtained with probes designed for ORFs of *S. scabies* and used for the analysis of *S. turgidiscabies* (see below).

© 2010 THE AUTHORS

The CR was highly similar in *S. scabies* ATCC49173 and most strains of *S. turgidiscabies*, including strains 32 and 323, whereas TR was clearly different from *S. scabies* in strains 32 and 323 (and strain Car8 used for probe design) (Fig. 2). The same conclusion was reached in microarray analysis using probes designed for the predicted ORFs of *S. scabies* (strain 87.22) (Fig. 3). Altogether, four divergent PAIs consisting of different variants of CR and TR could be detected in strains of *S. turgidiscabies*. CR and TR were 'scabies-like' in five strains (261, 304, 342, 368, 255), meaning that the probes designed for the ORFs of *S. scabies* 87.22 detected them and *S. scabies* ATCC48173 readily, whereas the probes designed for the PAI of *S. turgidiscabies* Car8 detected them poorly. Strains 32 and 323 had a 'scabies-like' CR, but a 'Car8-like' TR. Strain 287 was exceptional in that CR was not 'scabies-like' or 'Car8-like', indicated by the negative results with both types of probe, whereas TR was 'scabies-like' and 'Car8-like', indicated by the positive results with both types of probe. Strain 300 represented the fourth type of PAI variant, because CR and TR could not be detected with probes designed for either species. Because strain 300 caused common scab symptoms on potato tubers and a few genes of CR and TR were detected by PCR in this strain, it probably contains a PAI which is divergent from the strains of *S. turgidiscabies* and *S. scabies* used for probe design. Taken together, the data indicate that the horizontal transfer of PAI regions may occur among a wider range of plant-pathogenic *Streptomyces* species than known to date.

The data showed that *fas1* and *nos* in TR were undetectable or readily detectable, respectively, in strains whose other genes in TR showed opposite results in terms of identity with Car8. Therefore, these two loci of TR may undergo independent evolution. It is also noteworthy that the *fas* gene cluster sequences differ among strains of *S. turgidiscabies*, as interpreted from the data of this study. The *fas* genes in *S. turgidiscabies* are attributable to cytokinin production and gall formation in *Arabidopsis thaliana* and *Nicotiana tabacum* (Joshi and Loria, 2007), but their role in virulence on potato remains to be elucidated. Taken together, the species' origins of the fas operon (not detected in *S. scabies*) and the detected variants of CR and TR require further study.

Thaxtomin A is the key metabolite of *Streptomyces* spp. in their ability to cause common scab symptoms on potato tubers (Goyer *et al*., 1998; Lawrence *et al*., 1990; Loria *et al*., 2008). Therefore, PCR-based testing of txt genes has been proposed as a diagnostic means for the detection of common scab pathogens (Flores-Gonzáles *et al*., 2008; Lerat *et al*., 2009). However, this recommendation does not apply to *S. reticuliscabiei*, a strain of *S. turgidiscabies* which induces netted scab symptoms on potato tubers (Bouchek-Mechiche *et al*., 2000b, 2006) and, to our knowledge, does not produce thaxtomin. Our data also provided no evidence for the presence of the txt gene cluster or *txtR* required for the transcriptional induction of txt genes (Joshi *et al*., 2007a) in the type strain of *S. reticuliscabiei*. Furthermore, the data cast doubt on whether the detection of txt genes can be used generally for the detection of pathogenic strains of *S. turgidiscabies*. Strains 255 and 304 were unable to produce thaxtomin A and cause common scab symptoms under conditions inducive with other strains, but no systematic difference was detected between these and other strains of *S. turgidiscabies* in the content of txt or other PAI genes. Furthermore, the sequence variability of txt genes may cause challenges to probe and primer design. The use of virulence-related genes, such as *nec1* or *tom*, which contribute to virulence but are dispensable for pathogenicity (Joshi *et al*., 2007b; Seipke and Loria, 2008), has been found to be unreliable for the detection of plant-pathogenic *Streptomyces* strains in previous studies (Flores-Gonzáles *et al*., 2008; Kreuze *et al*., 1999; Wanner, 2006, 2009) and also this study. Probes for most CR genes generated low, if any, hybridization signals in strains 287 and 300, and only a few genes were detected by PCR. Failure to detect *nec1* in these two strains was also reported in a previous study (Kreuze *et al*., 1999). Nevertheless, these strains cause characteristic and severe symptoms of common scab on potato tubers (illustrated in Hiltunen *et al*., 2005; Kreuze *et al*., 1999).

Some data of this study were in accordance with the previously hypothesized horizontal transfer of PAI from *S. scabies* to *S. turgidiscabies* (Healy *et al*., 1999; Loria *et al*., 2006). For five strains of *S. turgidiscabies* (261, 304, 342, 386 and 255), the signals for genes in CR and TR were quite similar to those in *S. scabies* ATCC49173. Exchange of PAI regions between these two species is conceivable in Finland, where the species occur in the same potato fields and also the same scab lesions on tubers (Lehtonen *et al*., 2004). Sharing of the same pathogenic and virulence factors may provide *S. turgidiscabies* with a high potential to gradually replace *S. scabies* as the main common scab pathogen in areas in which both species occur, because, additionally, *S. turgidiscabies* is antagonistic to *S. scabies* (Hiltunen *et al*., 2009) and tolerates lower pH than *S. scabies* (Lindholm *et al*., 1997).

The hypothesis proposing the direct movement of the whole PAI from *S. scabies* to *S. turgidisacbies* (Healy *et al*., 1999; Loria *et al*., 2006) was challenged by a much wider distance between the putative CR and TR in *S. scabies* than in *S. turgidiscabies*. The putative genomic segments that might move horizontally between these species were mapped to 100-kb and 10-kb wide genomic segments of *S. scabies*, which correspond to CR and TR, respectively, and are highly similar to most strains of *S. turgidiscabies*, in contrast with the sequences flanking them. The putative CR and TR are *c*. 5 Mb apart in the genome sequence of *S. scabies* 87.22, as compared with only 0.3 Mb in *S. turgidiscabies* Car8 (Kers *et al*., 2005). If both regions were gained simultaneously from *S. scabies*, some chromosomal rearrangements or further transposition must have occurred later in *S. turgidisca-*

© 2010 THE AUTHORS
MOLECULAR PLANT PATHOLOGY © 2010 BSPP AND BLACKWELL PUBLISHING LTD

bies. Eight of nine PAI genes characterized in CR of *S. turgidiscabies* were among the 95 predicted ORFs in the 100-kb segment of *S. scabies*. It is plausible but not possible to confirm without additional genome sequence information whether all of these genes also exist in the PAI of *S. turgidiscabies*, because only *c*. 16 kb of the CR sequence is available (Kers *et al*., 2005). However, our data suggest that CR includes many additional as yet uncharacterized genes involved in pathogenicity and virulence. This may also be one reason for the unreliability encountered to date with the use of any particular currently known PAI gene or their combination as an indicator of plant pathogenicity in *Streptomyces* (Lerat *et al*., 2009; Wanner, 2009). Taken together, microarray analysis revealed and defined a large CR-related genomic region shared by *S. scabies* and *S. turgidiscabies*. This genomic segment might be mobile between the species and might modulate the virulence of the recipient strains, which requires further study.

Data from the microarray-based comparison of bacterial strains could also be used to infer genetic relatedness. Although phylogenetic analysis of 16S rRNA genes can be used for the reliable discrimination of more distantly related species, multilocus phylogeny is more efficient for resolving relationships among closely related species (Guo *et al*., 2008). Microarray analysis allows a genome-wide multilocus analysis of relatedness (Taboada *et al*., 2008), and has been applied to compare other *Streptomyces* species, but not those pathogenic on plants (Hsiao and Kirby, 2008, 2009). Our previous study showed that *S. turgidiscabies* and *S. scabies* can be distinguished readily from each other and also other *Streptomyces* species by microarray analysis using probes designed for genes of *S. scabies* (Aittamaa *et al*., 2008). In the present study, this approach was used to analyse the relatedness of strains of *S. turgidiscabies* with reference to the type strains of *S. scabies* and *S. reticuliscabiei*. Cluster analysis of microarray data was consistent with the results obtained by phylogenetic analysis of 16S rRNA gene sequences. However, importantly, microarray analysis revealed two strains of *S. turgidiscabies* that were isolated from the same geographical area and were the most closely related strains at the whole-genome level, but nevertheless contained clearly divergent regions of PAI, probably of different origins.

In conclusion, the results of this study provide evidence for multiple origins of pathogenicity- and virulence-related gene clusters in *S. turgidiscabies*, and show that PAI constitutes a mosaic of regions that may undergo independent evolution. The genome-wide approach for the comparison of bacterial genomes also provided new evidence on the genomic segments of *S. scabies* which may be horizontally transferred between species. The microarray approach can be used efficiently for the characterization of genetic variability between bacterial strains, and to carry out more detailed comparisons of the genomic areas of interest than obtained by previously used methods. Comprehensive comparison of *S. turgidiscabies* strains and the genomes of other plant-pathogenic *Streptomyces* species by microarray analysis can contribute significantly to the understanding of the genetic composition required for pathogenicity and virulence on plants, and opens up novel possibilities for diagnostic purposes. Arrays can be revised for more detailed analysis in the light of additional sequence data acquired by the ongoing genome sequencing of *Streptomyces* species.

EXPERIMENTAL PROCEDURES

Bacteria

The bacteria were originally isolated from potato scab lesions in Finland or northern Sweden (Table 1). Strains of *S. turgidiscabies* and *S. scabies* have been characterized for morphological and physiological characteristics (Lehtonen *et al*., 2004; Lindholm *et al*., 1997), and also 16S rRNA gene sequences (Kreuze *et al*., 1999), or identified by PCR using *S. scabies*- and *S. turgidiscabies*-specific primers designed for the 16S rRNA gene (Lehtonen *et al*., 2004). All strains were re-tested with these species-specific primers during this study. In addition, the 16S rRNA and intergenic regions were sequenced if sequences were not available (see below). The type strains of *S. scabies* (ATCC 49173) (Lambert and Loria, 1989a) and *S. reticuliscabiei* (DSM41804) (Bouchek-Mechiche *et al*., 2000a) were obtained from the respective type culture collections.

Bacteria were grown on potato dextrose agar (Sigma, Steinheim, Germany) at 28 °C, and plates were stored at 4 °C. Liquid cultures in potato dextrose broth or oat bran broth (OBB) (Sigma) were grown under shaking (120 rpm) at 28 °C, and cells were maintained at −70 °C in culture medium containing 20% glycerol (Kieser *et al*., 2000).

Pathogenicity tests

The pathogenicity of bacteria was tested in the glasshouse using large pots (20 L) or large boxes (Hiltunen *et al*., 2005; 2009) filled with a mixture of peat and sand (pH 6.0) (Kekkilä, Mellilä, Finland). The bacteria were grown in 200 mL of potato dextrose broth (Sigma) for 7 days. The volume of culture was increased to 1000 mL by the addition of sterile water and 500 mL of the suspension poured into two pots or boxes each. Soil was mixed thoroughly taking care not to contaminate bacteria between pots or boxes. After 2 days, two healthy potato minitubers of cv. Bintje or Amazon (Finnish Seed Potato Center, Tyrnävä, Finland) were planted into each pot or box. Potatoes were harvested and examined for scab lesions after 3 months.

© 2010 THE AUTHORS
 MOLECULAR PLANT PATHOLOGY © 2010 BSPP AND BLACKWELL PUBLISHING LTD

Thaxtomin A production

The bacteria were grown in thaxtomin A-inducing medium (OBB) (Goyer *et al*., 1998) at 28 °C for 7 and 14 days, followed by extraction of thaxtomins from the culture medium (400 mL) with an equal volume of ethyl acetate, as described by Hiltunen *et al*. (2006). The solvent phase was evaporated to a volume of 50 μL; 10 μL was analysed by TLC on silica gel 60 F_{254} plates (Merck GaA, Darmstadt, Germany) using chloroform with 10% methanol as a liquid phase. The samples dissolved in methanol were examined for thaxtomin A using HPLC, as described in Hiltunen *et al*. (2006), with a Waters 600E HPLC instrument and a Waters 991 photodiode array detector (Millipore Waters, Milford, MA, USA) using a Hypersil BDS C18 column (5 μm particle size, 4.6 × 150 mm; Agilent Technologies, Palo Alto, CA, USA). An acetonitrile concentration gradient from 20% to 50% at a flow rate of 0.8 mL/min for 20 min constituted the liquid phase. The retention time and shape of each peak detectable at a wavelength of 402 nm were compared with those of purified thaxtomin A used as a reference (Hiltunen *et al*., 2006) and also scanned at 200–500 nm.

Probe design for microarrays

The whole genome sequence of *S. scabies* strain 87.22 (FN554889) and the PAI sequences of *S. turgidiscabies* strain Car8 (accession nos. AY707079, AY707080, AY707081) (Kers *et al*., 2005) were used for probe design with OligoArray 2.1 software (Rouillard *et al*., 2003). The probes were designed for all predicted ORFs and their intergenic regions in the PAIs of *S. turgidiscabies* and 3919 randomly selected ORFs of *S. scabies* including the PAI regions. Probe specificity in OligoArray 2.1 is based on the BLAST algorithm and the estimation of the melting temperature (T_m). If T_m of hybridization between a candidate probe and a nonspecific target exceeds the given threshold, the probe is predicted to be nonspecific and not selected. Furthermore, those probe candidates that contain stable secondary structures and those containing a stretch of more than four identical nucleotides are not accepted. The target length of each probe was defined to be 40 nucleotides. Because of the high GC content of some sequences, the T_m range of probes was set at 82–90 °C.

A total of 113 different probes was designed for the ORFs and intergenic areas of PAIs in *S. turgidiscabies* (one to five probes per ORF), but no specific probe could be designed for the ORF stPAI25 (average T_m = 88.7 °C; standard deviation, 1.4). One probe was designed for each of the 3919 ORFs of *S. scabies* (average T_m = 88.0 °C; standard deviation, 1.2). A 20-nucleotide-long poly(T) tail was added to each probe. The probes were synthesized on an Agilent 8*15K custom array (Agilent, Santa Clara, CA, USA) (Wolber *et al*., 2006). Each of the eight subarrays synthesized on the same array contained the same probes. The probes for PAIs of *S. turgidiscabies* were synthesized in triplicate on each subarray.

DNA isolation and microarray hybridization

DNA was isolated from cultures grown in potato dextrose broth (Sigma) for 7 days. Cells were collected from 5 mL of culture, resuspended in 600 μL of TE buffer [50 mM Tris-HCl, pH 8, 20 mM ethylenediaminetetraacetic acid (EDTA)] and 200 μL of 5% sodium dodecylsulphate (VWR, Poole, UK), and 200 μL of proteinase K (1 mg/mL, Finnzymes, Espoo, Finland) was added. The mixture was incubated for up to 4 h at 37 °C until lysis. After 10 min of centrifugation at 9500 ***g*** the supernatant was extracted with equal volumes of phenol and chloroform, and the DNA was precipitated with two volumes of ethanol, washed with 70% ethanol, dried, dissolved in sterile water and diluted to a concentration of 50 ng/μL.

DNA samples (500 ng) were labelled with Cy3 or Cy5 dCTP (GE Healthcare, Amersham, Buckinghamshire, UK) and were purified using a BioPrime labelling kit (Invitrogen, Carlsbad, CA, USA). The DNA concentration and incorporation of the dye were checked with a Nanodrop analysator (Nanodrop Technologies, Wilmington, MA, USA) before and after labelling. Two DNA samples labelled with a different dye were hybridized on each subarray simultaneously. In the repeated experiment, the dyes were swapped between the samples. Hybridization was carried out at 65 °C for 18–20 h using the microarray manufacturer's reagents and an 8 × 15K CGH protocol (Agilent publication number G4427-90010; http://www.chem.agilent.com/).

Microarray data analysis

Microarray slides were scanned with a GenePix 4200 AL scanner (Axon Instruments, Foster City, CA, USA) using a pixel resolution of 5 μm. Image analysis and spot segmentation were performed with GenePix Pro 6.0 software. Spots were visually investigated and those showing anomalies were manually flagged and ignored in further analysis.

Spot intensity values were processed using the Limma package in R (R Development Core Team, 2004). Local background was subtracted from the foreground signal of each spot using the method 'normexp'. Offset 50 was added before transforming the signal values into the log2 domain. When two samples (each labelled with a different dye) were hybridized simultaneously on the same array to spare the numbers of microarrays needed, signals for both samples were analysed separately using single-channel measurements.

The 113 probes for the PAI of *S. turgidiscabies* were triplicated on the array. For signals of each triplicate set of probes, a median value of signals was calculated and, for the total set of 4032 probes (113 PAI and 3919 *S. scabies* probes), the array-specific

© 2010 THE AUTHORS

log2 intensity distributions were visually investigated. The shape of the signal distributions was bimodal, suggesting that the signals can be divided into two groups as in a previous study (Aittamaa *et al*., 2008), in which hybridizations with different bacterial species revealed that the peaks of high-intensity and low-intensity signals correspond to the target-specific and non-specific probes, respectively. Rather than determining the location of the smallest intensity value in the valley between the two peaks of kernel density representation (Aittamaa *et al*., 2008), the classification boundary dividing signals into two categories was based on an array-specific Gaussian mixture model decomposition of log2 signal intensities (Asyali *et al*., 2004).

The signals of the probes were converted into 1/0 values, corresponding to their location in the high or low signal peak, respectively, in the array-specific, two-component Gaussian mixture model. Parameters of the mixture model were computed in R using the quasi-Newton optimizer, as described in Venables and Ripley (2002). The classification boundary was calculated analytically for each mixture model by solving a second-order equation which gives two solutions, but only one is located between the centroids of the two mixture model components. This Bayesian classifier was used to binarize the signals.

After binarization of the probe intensities, the results from two independent hybridizations were averaged, which resulted in three possible values (0, 0.5 or 1) for each probe. Values 0 and 1 indicate consistent categorization in repeated tests. Probes were placed in order according to their locations in the PAI of *S. turgidiscabies* and the genome sequence of *S. scabies*, and signal values from different strains were visually investigated using TM4 MEV microarray analysis software (Saeed *et al*., 2003).

Genetic relatedness was inferred from microarray data using the *p* distance defined as the proportion of nucleotide sites which are different between two strains (Kumar *et al*., 2008). Microarray signal data were converted into a sequence of 1/0 signals (see above) and coded by nucleotides C and A, respectively, before importing them to MEGA (Kumar *et al*., 2008). Values between 1 and 0 were labelled as missing data.

The microarray data generated in this study, including log2 intensities and 1/0 categories, were deposited in Gene Expression Omnibus (GEO) according to Minimum Information About a Microarray Experiment (MIAME) and are available under accession number GSE18665.

Detection of conserved genomic regions using selected probe sequences

There were 298 probes designed for *S. scabies* ORFs which gave category 1 signals in the hybridization of all bacterial strains tested. These probes were used as a query in a local BLASTN search against genome sequences of *S. coelicolor* (NC_003888), *S. avermitilis* (BA000030) and *S. griseus* (AP009493), and 224 probes resulted in a BLAST hit with an *e*-value <1e-5. Among these probes, those which contained four nucleotide mismatches at most along the entire probe length of 40 nucleotides when compared with any of the target genomes were selected, which resulted in a final set of 120 probes. Gene annotations of the corresponding target sequences were collected from *S. coelicolor*, *S. avermitilis* and *S. griseus*, and are reported in Table S3.

PCR and sequencing

Primer pairs were designed for each ORF of the PAIs of *S. turgidiscabies* strain Car8 and also for amplification of the intergenic region between stPAI27 and stPAI28 (Table S2). Primers were also designed for the *nec1* gene, a nearby transposone pseudogene (tnp), their intergenic region, and *txtA*, *txtB* and *txtC* (Healy *et al*., 2000; 2002). PCR (25 μL) contained 0.2 μM of each primer, 0.16 mM deoxynucleoside triphosphates (dNTPs), 0 μL Dynazyme II polymerase in buffer containing 2 mM $MgCl_2$ (Finnzymes, Espoo, Finland) and 0.5 μL DNA. Universal primers pA and pH′ (Edwards *et al*., 1989) for the amplification of the 16S rRNA gene were used as a positive control. The PCR cycle was 92 °C for 2 min, followed by 35 cycles consisting of denaturation at 92 °C for 30 s, annealing at 55 or 55 °C for 30 s and extension for 1 min, and final extension at 72 °C for 5 min. PCR products were analysed by electrophoresis in a 1% agarose gel stained with ethidium bromide and visualized under UV light. Results were considered as positive when a product of the expected size was obtained with either annealing temperature.

Some of the PCR products were sequenced for verification. For sequencing of the products, a PCR volume of 50 μL was used. The 16S rRNA gene and the flanking intergenic regions were amplified as a single product with primers pA (Edwards *et al*., 1989) and L1r (Lane, 1991) as above, which resulted in a product of 1.9 kb. Nested primers 16Sf and 16Sr (Aittamaa *et al*., 2008) were used for sequencing the products at the Haartman Institute, University of Helsinki, Finland. The intergenic region alone was amplified by PCR from strains for which the 16S rRNA gene sequence was already available, and sequenced using the primers f1114 (Gutell and Fox, 1988) and L1r (Lane, 1991) (Table S2).

Neighbour-joining analysis of DNA sequences (Saitou and Nei, 1987) and drawing of phylogenetic trees were performed using MEGA (Kumar *et al*., 2008). Nucleotide sequences were aligned using CLUSTALW, and the Kimura two-parameter model (Kimura, 1980) was used for analysis.

ACKNOWLEDGEMENTS

We thank Maaret Suokas, Minna Alanen and Anu Kallinen for skilful technical assistance. Financial support from the Finnish Ministry of Agriculture and Forestry (grant 4541/501/2006) is gratefully acknowledged.

© 2010 THE AUTHORS
 MOLECULAR PLANT PATHOLOGY © 2010 BSPP AND BLACKWELL PUBLISHING LTD

REFERENCES

Aittamaa, M., Somervuo, P., Pirhonen, M., Mattinen, L., Nissinen, R., Auvinen, P. and Valkonen, J.P.T. (2008) Distinguishing bacterial pathogens of potato using a genome-wide microarray approach. *Mol. Plant Pathol.* **9**, 705–717.

Asyali, M., Shoukri, M., Demirkaya, O. and Khabar, K. (2004) Assessment of reliability of microarray data and estimation of signal thresholds using mixture modeling. *Nucleic Acids Res.* **32**, 2323–2335.

Bouchek-Mechiche, K., Gardan, L., Normand, P. and Jouan, B. (2000a) DNA relatedness among strains of Streptomyces pathogenic to potato in France: description of three new species, *S. europaeiscabiei* sp. nov. and *S. stelliscabiei* sp. nov. associated with common scab, and *S. reticuliscabiei* sp. nov. associated with netted scab. *Int. J. Syst. Evol. Microbiol.* **50**, 91–99.

Bouchek-Mechiche, K., Pasco, C., Andrivon, D. and Jouan, B. (2000b) Differences in host range, pathogenicity to potato cultivars and response to soil temperature among *Streptomyces* species causing common and netted scab in France. *Plant Pathol.* **49**, 3–10.

Bouchek-Mechiche, K., Gardan, L., Andrivon, D. and Normand, P. (2006) *Streptomyces turgidiscabies* and *Streptomyces reticuliscabiei*: one genomic species, two pathogenic groups. *Int. J. Syst. Evol. Microbiol.* **56**, 2771–2776.

Bukhalid, R.A., Chung, S.Y. and Loria, R. (1998) Nec1, a gene conferring a necrogenic phenotype, is conserved in plant-pathogenic *Streptomyces* spp. and linked to a transposase pseudogene. *Mol. Plant–Microbe Interact.* **11**, 960–967.

Bukhalid, R.A., Takeuchi, T., Labeda, D. and Loria, R. (2002) Horizontal transfer of the plant virulence gene, *nec1*, and flanking sequences among genetically distinct *Streptomyces* strains in the diastochromogenes cluster. *Appl. Environ. Microbiol.* **68**, 738–744.

Edwards, U., Rogall, T., Blöcker, H., Emde, M. and Böttger, E.C. (1989) Isolation and direct complete determination of entire genes. Characterization of a gene coding for 16S ribosomal RNA. *Nucleic Acids Res.* **17**, 7843–7853.

Flores-Gonzáles, R., Velasco, I. and Montes, F. (2008) Detection and characterization of Streptomyces causing potato common scab in Western Europe. *Plant Pathol.* **57**, 162–169.

Goyer, C., Faucher, E. and Beaulieu, C. (1996) *Streptomyces caviscabies* sp. nov. from deep-pitted lesions in potatoes in Quebec, Canada. *Int. J. Syst. Evol. Microbiol.* **46**, 635–639.

Goyer, C., Vachon, J. and Beaulieu, C. (1998) Pathogenicity of *Streptomyces scabies* mutants altered in thaxtomin A production. *Phytopathology*, **88**, 442–445.

Guo, Y., Zheng, W., Rong, X. and Huang, Y. (2008) A multilocus phylogeny of the *Streptomyces griseus* 16S rRNA gene cluster: use of multilocus sequence analysis for Streptomyces systematics. *Int. J. Syst. Evol. Microbiol.* **58**, 149–159.

Gutell, R.R. and Fox, G.E. (1988) A compilation of large subunit RNA sequences presented in a structural format. *Nucleic Acids Res.* **16** (Suppl), r175–r201.

Healy, F.G., Bukhalid, R.A. and Loria, R. (1999) Characterization of an insertion sequence element associated with genetically diverse plant pathogenic *Streptomyces* spp. *J. Bacteriol.* **181**, 1562–1568.

Healy, F.G., Wach, M., Krasnoff, S.B., Gibson, D.M. and Loria, R. (2000) The txtAB genes of the plant pathogen *Streptomyces acidiscabies* encode a peptide synthetase required for phytotoxin thaxtomin A production and pathogenicity. *Mol. Microbiol.* **38**, 794–804.

Healy, F.G., Krasnoff, S.B., Wach, M., Gibson, D.M. and Loria, R. (2002) Involvement of a cytochrome P450 monooxygenase in thaxtomin A biosynthesis by *Streptomyces acidiscabies*. *J. Bacteriol.* **184**, 2019–2029.

Hiltunen, L., Weckman, A., Ylhäinen, A., Rita, H., Richter, E. and Valkonen, J.P.T. (2005) Responses of potato cultivars to the common scab pathogens, *Streptomyces scabies* and *S. turgidiscabies*. *Ann. Appl. Biol.* **146**, 395–403.

Hiltunen, L., Laakso, I., Chobot, V., Hakala, K.S., Weckman, A. and Valkonen, J.P.T. (2006) Influence of thaxtomins in different combinations on growth of micropropagated potato shoot cultures. *J. Agric. Food Chem.* **54**, 3372 -3379.

Hiltunen, L., Ojanperä, T., Kortemaa, H., Richter, E., Lehtonen, M.J. and Valkonen, J.P.T. (2009) Interactions and biocontrol of pathogenic *Streptomyces* strains co-occurring in potato scab lesions. *J. Appl. Microbiol.* **106**, 199–212.

Hsiao, N. and Kirby, R. (2008) Comparative genomics of *Streptomyces avermitilis, Streptomyces cattleya, Streptomyces maritimus* and *Kitasatospora aureofaciens* using *Streptomyces coelicolor* microarray system. *Antonie van Leeuwenhoek* **93**, 1–25.

Hsiao, N. and Kirby, R. (2009) Two-component signal transduction systems in *Streptomyces* and related organisms studied using DNA comparative microarray analysis. *Antonie van Leeuwenhoek* **95**, 189–206.

Johnson, E.G., Joshi, M.V., Gibson, D.M. and Loria, M. (2007) Cello-oligosaccharides released from host plants induce pathogenicity in scab-causing *Streptomyces* species. *Physiol. Mol. Plant Pathol.* **71**, 18–25.

Johnson, E.G., Sparka, J.P., Dzikovski, B., Crane, B.R., Gibson, D.M. and Loria, R. (2008) Plant-pathogenic *Streptomyces* species produce nitric oxide synthase-derived nitric oxide in response to host signals. *Chem. Biol.* **15**, 43–50.

Joshi, M. and Loria, R. (2007) *Streptomyces turgidiscabies* possesses a functional cytokinin biosynthetic pathway and produces leafy galls. *Mol. Plant–Microbe Interact.* **20**, 751–758.

Joshi, M.V., Bignell, D.R., Johnson, E.G., Gibson, D.M. and Loria, R. (2007a) The AraC/XylS regulator TxtR modulates thaxtomin biosynthesis and virulence on *Streptomyces scabies*. *Mol. Microbiol.* **66**, 633–642.

Joshi, M., Rong, X., Moll, S., Kers, J., Franco, C. and Loria, R. (2007b) *Streptomyces turgidiscabies* secretes a novel virulence protein, Nec1, which facilitates infection. *Mol. Plant–Microbe Interact.* **20**, 599–608.

Kers, J.A., Cameron, C.D., Joshi, M.V., Burkhalid, R.A., Morello, J.E., Wach, M.J., Gibson, D.M. and Loria, R. (2005) A large, mobile pathogenity island confers plant pathogenicity on *Streptomyces* species. *Mol. Microbiol.* **55**, 1025–1033.

Kieser, T., Bibb, M.J., Buttner, M.J., Chater, K.F. and Hopwood, D.A. (2000) *Practical Streptomyces Genetics*. Norwich: The John Innes Foundation.

Kimura, M. (1980) A simple method for estimating evolutionary rates of base substitutions through comparative studies of nucleotide sequences. *J. Mol. Evol.* **16**, 111–120.

Kreuze, J.F., Suomalainen, S., Paulin, L. and Valkonen, J.P.T. (1999) Phylogenetic analysis of 16S rRNA genes and PCR analysis of the *nec1* gene from *Streptomyces* spp. causing common scab, pitted scab and netted scab in Finland. *Phytopathology*, **89**, 462–469.

© 2010 THE AUTHORS
MOLECULAR PLANT PATHOLOGY © 2010 BSPP AND BLACKWELL PUBLISHING LTD

Kumar, S., Dudley, J., Nei, M. and Tamura, K. (2008) MEGA: a biologist-centric software for evolutionary analysis of DNA and protein sequences. *Brief. Bioinform.* **9**, 299–306.

Lambert, D.H. and Loria, R. (1989a) *Streptomyces scabies* sp. nov., nom. rev. *Int. J. Syst. Bacteriol.* **39**, 387–392.

Lambert, D.H. and Loria, R. (1989b) *Streptomyces acidiscabies* sp. nov. *Int. J. Syst. Bacteriol.* **39**, 393–396.

Lane, D. (1991) 16S/23S sequencing. In: *Nucleic Acid Techniques in Bacterial Systematics* (Stackebrandt, E. and Goodfellow, M., eds), pp. 115–175. Chichester: John Wiley & Sons.

Lawrence, C.H., Clark, M.C. and King, R.R. (1990) Induction of common scab symptoms in aseptically cultured potato tubers by the vivotoxin, thaxtomin. *Phytopathology*, **80**, 606–608.

Lehtonen, M.J., Rantala, H., Kreuze, J.F., Bång, H., Kuisma, L., Koski, P., Virtanen, E., Vihlman, K. and Valkonen, J.P.T. (2004) Occurrence and survival of potato scab pathogens (*Streptomyces* species) on tuber lesions: quick diagnosis based on a PCR-based assay. *Plant Pathol.* **53**, 280–287.

Lerat, S., Simao-Beaunoir, A.M. and Beaulieu, C. (2009) Genetic and physiological determinants of *Streptomyces scabies* pathogenicity. *Mol. Plant Pathol.* **10**, 579–585.

Lindholm, P., Kortemaa, H., Kokkola, M., Haahtela, K., Salkinoja-Salonen, M. and Valkonen, J.P.T. (1997) *Streptomyces* spp. isolated from potato scab lesions under Nordic conditions in Finland. *Plant Dis.* **81**, 1317–1322.

Loria, R., Kers, J. and Joshi, M. (2006) Evolution of plant pathogenicity in Streptomyces. *Annu. Rev. Phytopathol.* **44**, 469–487.

Loria, R., Bignell, D.R.D., Moll, S., Huguet-Tapia, J., Joshi, M.V., Johson, E.G., Seipke, R.F. and Gibson, D.M. (2008) Thaxtomin biosynthesis: the path to plant pathogenicity in the genus *Streptomyces*. *Antonie van Leeuwenhoek* **94**, 3–10.

Miyajima, K., Tanaka, F., Takeuchi, T. and Kuninaga, S. (1998) *Streptomyces turgidiscabies* sp. nov. *Int. J. Syst. Bacteriol.* **48**, 495–502.

Naito, S., Kondo, N., Akino, S., Ogoshi, A. and Tanaka, F. (eds) (2004) *Novel Approaches to Control of Potato Scab*. Proc. Int. Potato Scab Symp. (IPSS 2004). Sapporo: Hokkaido University.

Park, D.H., Kim, J.S., Kwon, S.W., Wilson, C., Yu, Y.M., Hur, J.H. and Lim, C.K. (2003) *Streptomyces luridiscabies* sp. nov., *Streptomyces puniciscabies* sp. nov. and *Streptomyces niveiscabies* sp. nov., which cause potato common scab disease in Korea. *Int. J. Syst. Evol. Microbiol.* **53**, 2049–2054.

R Development Core Team (2004) *R: A Language and Environment for Statistical Computing*. Vienna: R Foundation for Statistical Computing. Available from URL: http://www.R-project.org

Rouillard, J., Zuker, M. and Gulari, E. (2003) OligoArray 2.0: design of oligonucleotide probes for DNA microarrays using a thermodynamic approach. *Nucleic Acids Res.* **31**, 3057–3062.

Saeed, A.I., Sharov, V., White, J., Li, J., Liang, W., Bhagabati, N., Braisted, J., Klapa, M., Currier, T., Thiagarajan, M., Sturn, A., Snuffin, M., Rezantsev, A., Popov, D., Ryltsov, A., Kostukovich, E., Borisovsky, I., Liu, Z., Vinsavich, A., Trush, V. and Quackenbush, J. (2003) TM4: a free, open-source system for microarray data management and analysis. *Biotechniques*, **34**, 374–378.

Saitou, N. and Nei, M. (1987) The neighbor-joining method: a new method for reconstructing phylogenetic trees. *Mol. Biol. Evol.* **4**, 406–425.

Scheible, W.R., Fry, B., Kochevenko, A., Schindelasch, D., Zimmerli, L., Somerville, S., Loria, R. and Somerville, C.R. (2003) An Arabidopsis mutant resistant to thaxtomin A, a cellulose synthesis inhibitor from *Streptomyces scabies*. *Plant Cell*, **15**, 1781–1794.

Seipke, R.F. and Loria, R. (2008) *Streptomyces scabies* 87-22 possesses a functional tomatinase. *J. Bacteriol.* **190**, 7684–7692.

Taboada, E.N., MacKinnon, J.M., Luebbert, C.C., Gannon, V.P.J., Nash, J.H.E. and Rahn, K. (2008) Comparative genomic assessment of Multi-Locus Sequence Typing: rapid accumulation of genomic heterogeneity among clonal isolates of *Campylobacter jejuni*. *BMC Evol. Biol.* **8**, 229. doi:10.1186/1471-2148-8-229.

Takeuchi, T., Sawada, H., Tanaka, F. and Matsuda, I. (1996) Phylogenetic analysis of *Streptomyces* spp. causing potato scab based on 16s rRNA sequences. *Int. J. Syst. Bacteriol.* **46**, 476–479.

Toth, L., Akino, S., Kobayashi, K., Doi, A., Tanaka, F. and Ogoshi, A. (1998) Production of thaxtomin A by *Streptomyces turgidiscabies*. *Soil Microorg.* **51**, 29–34.

Venables, W. and Ripley, B. (2002) *Modern Applied Statistics with S*, 4th edn. New York: Springer-Verlag.

Wanner, L. (2006) A survey of genetic variation in *Streptomyces* isolates causing potato common scab in the United States. *Phytopathology*, **96**, 1363–1371.

Wanner, L. (2009) A patchwork of Streptomyces species isolated from potato common scab lesions in North America. *Am. J. Potato Res.* **86**, 247–264 (doi 10.1007/s12230-009-9078-y).

Wolber, P.K., Collins, P.J., Lucas, A.B., De Witte, A. and Shannon, K.W. (2006) The Agilent in situ-synthesized microarray platform. *Methods Enzymol.* **410**, 28–57.

SUPPORTING INFORMATION

Additional Supporting Information may be found in the online version of this article:

Fig. S1 Detection of production of thaxtomin A by thin layer chromatography.

Table S1 Polymerase chain reaction (PCR) primers used for PCR and sequencing.

Table S2 Detection of open reading frames in the pathogenicity island of strains of *Streptomyces turgidiscabies* and the type strains of *S. scabies* (ATCC49173) and *S. reticuliscabiei* (DSM41804) by polymerase chain reaction.

Table S3 A total of 120 probes designed for open reading frames (ORFs) of *Streptomyces scabies* which gave category 1 signals to all strains of *S. scabies* and *S. turgidiscabies* tested and, furthermore, whose sequences are >90% identical to ORF sequences of *S. coelicolor*, *S. avermitilis* and *S. griseus*. (An Excel file, not embedded in this manuscript document)

Please note: Wiley-Blackwell are not responsible for the content or functionality of any supporting materials supplied by the authors. Any queries (other than missing material) should be directed to the corresponding author for the article.

© 2010 THE AUTHORS
 MOLECULAR PLANT PATHOLOGY © 2010 BSPP AND BLACKWELL PUBLISHING LTD