# SCIENTIFIC REPORTS

Received: 18 March 2015 Accepted: 09 June 2015 Published: 07 July 2015

## **OPEN** Functional characterization of the principal sigma factor RpoD of phytoplasmas viα an in vitro transcription assay

Chihiro Miura<sup>1</sup>, Ken Komatsu<sup>2</sup>, Kensaku Maejima<sup>1</sup>, Takamichi Nijo<sup>1</sup>, Yugo Kitazawa<sup>1</sup>, Tatsuya Tomomitsu<sup>1</sup>, Akira Yusa<sup>1</sup>, Misako Himeno<sup>1</sup>, Kenro Oshima<sup>1</sup> & Shigetou Namba<sup>1</sup>

Phytoplasmas (class, *Mollicutes*) are insect-transmissible and plant-pathogenic bacteria that multiply intracellularly in both plants and insects through host switching. Our previous study revealed that phytoplasmal sigma factor rpoD of OY-M strain ( $rpoD_{OY}$ ) could be a key regulator of host switching, because the expression level of  $rpoD_{\alpha\gamma}$  was higher in insect hosts than in plant hosts. In this study, we developed an in vitro transcription assay system to identify RpoD<sub>OV</sub>-dependent genes and the consensus promoter elements. The assay revealed that  $RpoD_{0Y}$  regulated some housekeeping, virulence, and host-phytoplasma interaction genes of OY-M strain. The upstream region of the transcription start sites of these genes contained conserved -35 and -10 promoter sequences, which were similar to the typical bacterial RpoD-dependent promoter elements, while the -35 promoter elements were variable. In addition, we searched putative RpoD-dependent genes based on these promoter elements on the whole genome sequence of phytoplasmas using in silico tools. The phytoplasmal RpoD seems to mediate the transcription of not only many housekeeping genes as the principal sigma factor, but also the virulence- and host-phytoplasma interaction-related genes exhibiting host-specific expression patterns. These results indicate that more complex mechanisms exist than previously thought regarding gene regulation enabling phytoplasmas to switch hosts.

In the regulation of bacterial gene expression, the initiation of transcription, mediated by a DNA-dependent RNA polymerase (RNAP) holoenzyme, plays an important role as the first step in the regulation process. The RNAP holoenzyme consists of a core enzyme (subunit composition  $\alpha 2\beta\beta'\omega$ ) with catalytic activity of RNA polymerization, and an additional subunit known as a sigma factor involved in promoter recognition and DNA melting<sup>1</sup>. Most bacteria have multiple sigma factors that recognize different sets of promoters as key regulators of stress responses to environmental changes and basal gene expression. RpoD (also known as sigma 70) is the principal (primary) and well-studied sigma factor responsible for the transcription of housekeeping genes in most bacteria<sup>1</sup>. The intracellular concentration of RpoD in Escherichia coli is maintained at a constant level under various growth conditions<sup>2</sup>. The sigma 70 family proteins, including RpoD, contain four conserved regions designated as 1 to 4<sup>3</sup>. In general, two DNA binding domains that reside in regions 2 and 4 recognize conserved promoter hexamer sequences (promoter elements) around positions approximately 10 and 35 nucleotides upstream, respectively, of the transcription start sites (TSSs)<sup>1</sup>. Previous studies have revealed that in many bacteria, RpoD interacts with the two archetypal promoter elements (-35 5'-TTGACA-3' and -10 5'-TATAAT-3' separated by an about 17-bp spacer)<sup>4</sup>.

<sup>1</sup>Graduate School of Agricultural and Life Sciences, The University of Tokyo, 1-1-1 Yayoi, Bunkyo-ku, Tokyo, 113-8657, Japan. <sup>2</sup>Graduate School of Agriculture, Tokyo University of Agriculture and Technology, 3-5-8 Saiwaicho, Fuchu, Tokyo 183-8509, Japan. Correspondence and requests for materials should be addressed to S.N. (email: anamba@mail.ecc.u-tokyo.ac.jp)

Phytoplasmas (class *Mollicutes*, genus '*Candidatus* Phytoplasma spp.') infect hundreds of plant species and cause devastating yield losses in various crops worldwide<sup>5,6</sup>. Phytoplasmas are biologically unique in that they can parasitize both plants (Kingdom Plantae) and insects (Kingdom Animalia). Although studies have shown that some phytoplasmas are vertically transmitted at low rates to plant embryos or insect progeny, they depend completely on "host switching" between plants and insects for their survival and dispersal<sup>7,8</sup>. Phytoplasmas dramatically alter their gene expression in response to host switching. For example, *tengu* and *phyl1*, phytoplasma genes encoding a secreted protein, are more highly expressed in plant hosts than insect hosts, and respectively induce phytoplasma-specific symptoms in plants such as witches' broom and phyllody<sup>9,10</sup>. Moreover, we previously showed that at least 33% of the genes in the genome of '*Ca*. P. asteris' onion yellows strain (OY-M) are differentially expressed when grown in plant versus insect hosts<sup>11</sup>.

Although the alternation of gene expression in response to host switching is assumed to be important for the host adaptation of phytoplasmas, how they regulate their gene expression to adapt to the two distinct intracellular environments (i.e., plant and insect cells) is unclear. To date, only limited studies have addressed the gene regulatory mechanisms in phytoplasmas due to the difficulties associated with their genetic engineering. All five genome-sequenced strains of phytoplasma have two types of sigma factors, RpoD and FliA<sup>12-16</sup>. Phytoplasmal RpoD, which has high amino acid sequence similarity with RpoD of E. coli, is highly conserved in each of the phytoplasma genomes as a single-copy gene (see Supplementary Fig. S1 online). Phytoplasmal RpoD contains all four conserved regions typical for sigma 70-type sigma factors. An amino acid sequence alignment of the RpoD homologous proteins shows that their C-terminal half, which contains regions 2 to 4 largely conserved in sigma factors, is almost identical to each other, whereas the N-terminal half containing region 1 was less conserved (see Supplementary Fig. S2 online). Given that regions 2 and 4 of RpoD are involved in the recognition of -10 and -35 promoter elements, respectively<sup>1</sup>, RpoDs of each phytoplasma likely have similar promoter recognition specificity. However, phytoplasmal *fliA* has detectable sequence similarity with an alternative sigma factor sigma 28 of E. coli or the extracytoplasmic function (ECF) sigma factor subfamily and is present in each of the phytoplasma genomes as a multi-copy gene, with the exception of 'Ca. P. mali'; these FliAs encode proteins containing only region 2, or regions 2 and 4 (see Supplementary Fig. S3 online)<sup>12-16</sup>.

The genome of OY-M contains 13 copies of putative sigma factor genes, one of which is categorized as rpoD (PAM\_628;  $rpoD_{OY}$ ) and the others are categorized as fliA ( $fliA_{OY}$ ). Despite the role of RpoD in E. coli and many other bacteria in regulating housekeeping genes<sup>17,18</sup>, the expression level of rpoD of OY-M ( $rpoD_{OY}$ ) is 4.0 times higher in insect hosts than in plant hosts<sup>11,19</sup>, even though actual accumulation level of RpoD<sub>OY</sub> protein is unknown. These data raise doubts as to whether RpoD<sub>OY</sub> functions as the principal sigma factor. A few studies have approached this issue using ex vivo or in silico methods. A recent study using an E. coli-based ex vivo reporter assay (ECERA) system that evaluates the interaction between phytoplasmal promoters and sigma factors based on the luciferase reporter activity in E. coli cells showed that RpoD<sub>OY</sub> activates the promoter regions (approximately 400 bp) of several housekeeping genes highly expressed in insect hosts<sup>19</sup>. These results suggest that RpoD<sub>OY</sub> has the potential to function as the principal sigma factor and plays a major role in infected insect hosts. However, due to possible indirect effects of cognate transcription factors in living *E. coli* cells, avoiding false positives and negatives in this EcERA system was difficult. Thus, this system does not appear to be suitable for further research to identify promoter elements recognized by RpoD<sub>OY</sub>. In another study, bioinformatic prediction using E. coli RpoD-dependent promoter elements was performed to identify phytoplasma RpoD-dependent promoters. However, obtaining precise estimates using this approach is difficult, since phytoplasma genomes, as well as *E. coli* RpoD-dependent promoter elements, are strongly AT-biased<sup>4,20</sup>. These results indicate that novel approaches are needed to provide more accurate measurements of the promoter activity regulated by RpoD<sub>OY</sub> to determine its function.

In this study, to elucidate the role of  $\text{RpoD}_{OY}$  in phytoplasma gene expression in detail, we developed an *in vitro* transcription assay of phytoplasma genes and analyzed their promoter activity using this assay. The consensus  $\text{RpoD}_{OY}$ -dependent promoter elements were identified for the first time. The genome-wide promoter prediction using this newly identified essential promoter elements revealed at least 88 genes that were regulated by  $\text{RpoD}_{OY}$ , which was involved in the transcription of not only many housekeeping genes, but also virulence- and host-phytoplasma interaction-related genes. We discuss the mechanism of transcription regulation associated with host switching between plants and insects based on the function of  $\text{RpoD}_{OY}$ .

#### Results

**RpoD**<sub>OY</sub> **recognizes two** *rrnB* **promoters.** RpoD recognizes specific promoter elements located at positions 35 and 10 bp upstream of  $TSSs^1$ , and hence, an experimental determination of TSSs provides important clues in estimating their upstream promoter elements. For the purpose of identifying promoter elements recognized by  $RpoD_{OY}$ , we investigated the TSSs of the 16S ribosomal RNA (*rrn*) B gene of OY-M, as *rrn* is often transcribed from RpoD-dependent promoter elements in many bacteria<sup>21-24</sup>. A 5' rapid amplification of cDNA ends (5' RACE) analysis of *rrnB* was performed using total RNA from phytoplasma-infected plants, resulting in the detection of three TSSs located 91, 254, and 444 bp upstream of *rrnB* (Fig. 1a). The same results were obtained from total RNA extracted from phytoplasma-infected insects. We designated these TSSs as P1, P2, and P3, respectively, and estimated their putative promoter



Figure 1. Identification of *rrnB* TSSs and analysis of *rrnB* promoter activity based on the *in vitro* transcription assay. (a) Schematic representation of the upstream promoter region and TSSs of the *rrnB* gene. The *rrnB* TSSs identified by 5' RACE analysis, designated as P1, P2, and P3 (91, 254, 444 nt upstream of *rrnB*, respectively), are represented by arrows. Putative –35 and –10 promoter elements of each of the three TSSs are underlined with dotted and continuous lines, respectively. (b) *In vitro* transcription assays using the RNAP holoenzyme with RpoD<sub>OY</sub>. RNAP<sub>Ec</sub> and a DNA template were incubated with NTP, including [ $\gamma$ -32P]CTP in the absence (–) or presence (+) of RpoD<sub>OY</sub>. A 784-bp DNA fragment named *PrrnB* covering the region from –500 to +284 of *rrnB* was used as a template. White arrowheads indicate the positions of the transcripts that are possibly transcribed from P2 and P3.

.....

elements as follows: P1 promoter elements, -35 5'-TTCACA-3' and -10 5'-TAATCT-3'; P2 promoter elements, -35 5'-TTGCTA-3' and -10 5'-TATAAT-3'; and P3 promoter elements, -35 5'-TTGCCA-3' and -10 5'-TATAAT-3'. Among these promoter elements, putative P2 and P3 promoter elements were highly similar to each other.

To identify which of these putative rrnB promoter elements are recognized by RpoD<sub>OV</sub>, we performed in vitro transcription assays, which have been used in studies of bacterial transcription systems. Some examples of the assays include a study identifying core promoter elements in a given input DNA sequence and then trying to correlate these elements to DNA-binding proteins such as sigma factors<sup>25</sup>. For in vitro transcription assays using RpoD<sub>OY</sub>, we reconstituted the RNAP holoenzyme with a commercially available E. coli RNAP core enzyme (RNAP<sub>Ec</sub>) and purified recombinant RpoD<sub>OY</sub> (RNAP<sub>Ec</sub>-RpoD<sub>OY</sub>; see Supplementary Fig. S4 online). A 784-bp DNA fragment from -500 to +284 of rrnB, named PrrnB and covering the three identified TSSs (P1, P2, and P3) and their putative promoter regions, was used as a template for transcription reactions. As shown in Fig. 1b, two major transcripts of approximately 500 and 700 nt were observed in an RpoD<sub>OY</sub>-dependent manner. The sizes of these transcripts corresponded to the expected sizes of transcripts initiated from P2 (254 bp upstream of rrnB) and P3 (444 bp upstream of rrnB), respectively. No transcripts corresponding to that transcribed from P1 (91 bp upstream of rrnB), which was expected to be  $375 \, \text{nt}$ , were detected. These results indicate that  $\text{RpoD}_{OY}$  is compatible with the heterologous E. coli RNAP to initiate transcription as the RNAP<sub>Ec</sub>-RpoD<sub>OY</sub> holoenzyme. Our results also demonstrated that the RNAP holoenzyme with RpoD<sub>OY</sub> recognizes the rrnB promoter elements upstream of P2 and P3, but not P1.

Identification of the *rrnB* core promoter sequence recognized by RpoD<sub>OY</sub>. To define the core promoter elements of rrnB, we focused on the putative P2 promoter and performed in vitro transcription assays. We introduced a series of double-base substitutions to GG into the putative P2 -35 and -10 promoter elements of a DNA fragment containing -400 to -1 of rrnB, named PrrnB\_P2 (Fig. 2a). In the putative -35 promoter element of P2, the double-base substitution of TT to GG at positions -35 and -34 (mt2) resulted in a drastic reduction of the transcript level to 16% compared with the intact PrrnB P2 (Fig. 2b), suggesting decreases in activity of the putative P2 promoter due to these mutations. In contrast, nucleotide substitutions at positions -37 and -36 (mt1), or -33 to -28 (mt3, mt4, and mt5), did not alter promoter activity. Similar results were obtained with the substitutions of TT to AA at the same positions (Fig. 2c). These results suggest that TT at positions -35 and -34 is crucial for promoter activity mediating transcription from rrnB P2. In the P2 -10 promoter element, the substitutions at positions from -12 to -7 (mt12, mt13, and mt14) also resulted in a reduction in promoter activity to 6-22% compared with the intact PrrnB\_P2 (Fig. 2d). Substitutions at positions -14 and -13 (mt11), and -6 and -5 (mt15), of P2 slightly decreased promoter activity (Fig. 2d). These results suggest that the conserved hexamer, 5'-TATAAT-3', is crucial for promoter activity. These findings indicate that at a minimum, the upstream sequence 5'-TT-21bp-TATAAT-3' of rrnB P2 is essential for recognition by RpoD<sub>OY</sub>. This is supported by the fact that the same sequence exists in the putative P3 promoter elements (-35 5'-TTGCCA-3' and -10 5'-TATAAT-3'), but not in the putative P1 promoter elements (-35 5'-TTCACA-3' and -10 5'-TAATCT-3').

RNAP holoenzyme containing RpoDoy mediates the transcription of various categories of genes. To investigate the OY-M genes regulated by RpoD<sub>OY</sub> we performed *in vitro* transcription assays using other templates. Many other bacteria housekeeping genes have sigma 70-type promoters<sup>26-29</sup>, so we used the upstream regions of four housekeeping genes of OY-M [the protein chain initiation factor IF-3 (infC), 50S ribosomal subunit protein L13 (rplM), 30S ribosomal subunit protein S4 (rpsD), and RNA polymerase sigma70 factor (rpoD) genes] for in vitro transcription as templates (PinfC, PrplM, PrpsD, and PrpoD, respectively). In addition, the upstream region of the molecular chaperone gene (*ibpA*) was used as a template (PibpA) because a previous study suggested that it is regulated by RpoD in an AT-rich bacterium 'Ca. Blochmannia floridanus'30, while in some other bacteria, *ibpA* is reported to be regulated by an alternative heat shock sigma factor RpoH<sup>31,32</sup>. Our in vitro transcription assays revealed that RNAP<sub>Ec</sub>-RpoD<sub>OY</sub> produced specific transcripts from the four templates (PinfC, PrplM, PrpsD, and PibpA), but did not produce a transcript from PrpoD (Fig. 3). When only RNAP<sub>Ec</sub>, which lacks RpoD<sub>OV</sub>, was added to the reaction, no specific transcripts were observed from any of these templates (Fig. 3). These results indicate that RpoD<sub>OY</sub> mediates the transcription of many phytoplasma housekeeping genes, although with some exceptions. Considering the previous findings that RNAP containing the principal sigma factor transcribes the majority of the housekeeping genes<sup>3</sup>,  $RpoD_{OY}$  is likely to play a role as the principal sigma factor.

To examine whether  $RNAP_{Ec}$ -RpoD<sub>OY</sub> recognizes OY-M gene promoters other than housekeeping genes, we performed *in vitro* transcription assays using five additional templates (PPAM157, PPAM289, PPAM486, Ptengu, and Pamp) containing upstream regions of genes associated with virulence or host–phytoplasma interactions: PAM157 (putative secreted protein), PAM289 (adhesin-like protein)<sup>33</sup>, PAM486 (putative secreted protein), tengu (secreted and virulence-related protein)<sup>9</sup>, and amp (insect transmissibility-related protein)<sup>34</sup>. In previous studies, we reported that PAM157 and PAM289 were highly expressed in insect hosts and PAM486 and tengu were highly expressed in plant hosts<sup>9,11</sup>, while amp was expressed at comparable levels in both plant and insect hosts<sup>11</sup>. The *in vitro* transcription assays revealed that RNAP<sub>Ec</sub>–RpoD<sub>OY</sub> produced specific transcripts from all templates, PPAM157, PPAM289, PPAM486, Ptengu, and Pamp (Fig. 3). These results suggest that an RNAP holoenzyme containing RpoD<sub>OY</sub> recognizes not only housekeeping genes, but also the genes related to virulence and host–phy-toplasma interactions.

Identification of consensus RpoD<sub>ov</sub>-dependent promoter elements. To more closely define the RpoD<sub>OY</sub>-dependent promoter elements, we performed a 5' RACE analysis of nine genes shown in this study to be transcribed by RNAP<sub>Ec</sub>-RpoD<sub>OY</sub> (Fig. 3; *ibpA*, *infC*, *rplM*, *rpsD*, *PAM157*, *PAM289*, *PAM486*, tengu, and amp) using total RNA from phytoplasma-infected plants or insects. As a result, we mapped the 5'-ends of ibpA, infC, rplM, rpsD, PAM157, PAM289, PAM486, tengu, and amp transcripts at 170, 415, 166, 301, 177, 222, 52, 153, and 81 nt upstream, respectively, of their start codons (Fig. 4a). In agreement with a previous study on the detection and identification of mycoplasma promoter sequences<sup>35</sup>, the 5'-end of these transcripts was either adenine or guanine. The predicted size from these identified TSSs to the 3'-ends of the above-mentioned in vitro transcription templates roughly corresponded to the length of transcripts produced by the *in vitro* transcription assays (Fig. 3). Next, to further characterize the RpoD<sub>OY</sub>-dependent promoter elements, we searched for consensus promoter elements using a motif finding tool, BioProspector, and found conserved -35 and -10 hexamers located at appropriate positions upstream of the TSSs (Fig. 4a). The consensus -10 promoter element (5'-TAtAAT-3') was found in all sequences examined (Fig. 4b). The consensus -35 promoter element (5'-TTgaca-3') was also found, even though this element was less conserved compared with the -10 promoter element (Fig. 4b). The spacing between the -35 and -10 promoter elements could vary from 17 to 19 nt. We also found two relatively











conserved regions, an 'extended -10 motif' (5'-TnTG-3') positioned around -17 to -14 and an A-rich region positioned around -42 to -39 (see Supplementary Fig. S5 online), which are common features of the promoter regions in other bacteria<sup>36,37</sup>.

Genome-wide prediction of RpoD<sub>oy</sub>-dependent genes. Based on the consensus RpoD<sub>OY</sub>-dependent promoter elements identified in this study with several sequence variants (Fig. 4b; [TC][AT][GC][AC][TC][AT]N<sub>17-19</sub>TA[AT]AA[AT]), we searched putative RpoD<sub>OY</sub>-dependent genes that possess these promoter elements on the whole genome sequence of the OY-M phytoplasma<sup>12</sup> using a DNA motif-finding program, RSA-tools facilities. Among a total of 540 sequence hits, we found 103 putative RpoD<sub>OV</sub>-dependent promoter elements located within a 500-bp upstream region of either ATG initiation codons or the 5'-ends of mature tRNA and rRNA of OY-M genes, which could mediate the transcription of at least 88 genes (about 12% of all OY-M genes; see Supplementary Table S3 online). Among these 103 putative promoters, twenty-two promoters (21.4%) contained a 5'-TG-3' motif at the extended -10 region, including seventeen promoters (16.5%) with the 5'-TnTG-3' motif. These results agree with a previous study of E. coli promoters<sup>38</sup>. Subsequently, these 88 putative RpoD<sub>OY</sub>-dependent genes were classified into Clusters of Orthologous Groups<sup>39</sup> based on their predicted functions, which showed that 25 genes (29%) belong to the category of information storage (replication, transcription, and translation), and 11 genes (12%) and 8 genes (9%) belong to the category of metabolism



**Figure 4. Sequence conservation in the RpoD**<sub>OY</sub>-dependent promoter region. (a) Putative promoter sequences deduced from the alignment of upstream sequences of TSSs identified by 5' RACE analysis. Bold letters indicate the putative -35 and -10 promoter elements. '+1' represents the position of the TSS. Numbers of the right side show the distance to the 5'-end of mature rRNA or ATG. (b) Consensus sequences of RpoD<sub>OY</sub>-dependent promoters. The consensus sequences of -35 (left) and -10 (right) promoter elements were determined using the BioProspector program<sup>63</sup> and illustrated with the WebLogo tool<sup>64</sup>.

(e.g., ABC-transporters) and cellular processes (e.g., co-chaperonin and zinc proteases), respectively (see Supplementary Fig. S6 online). Previous studies have suggested that phytoplasma genes that are highly expressed in insect hosts are likely to be regulated by  $\text{RpoD}_{OY}$ , which is also highly expressed in insect rather than plant hosts<sup>11,19</sup>. However, we did not find a correlation between the  $\text{RpoD}_{OY}$  dependence of phytoplasma genes predicted in this study and their host-specific expression pattern during host switching as described in our previous microarray analysis<sup>11</sup>. According to the microarray results, among the 88 putative  $\text{RpoD}_{OY}$ -dependent genes, as many as 32 genes showed no significant differences in expression pattern between plant and insect hosts, while 21 and 10 genes were upregulated more than twofold in plant and insect hosts, respectively.

To examine whether RpoD homologous proteins in other species of phytoplasma also regulate similar sets of genes such as OY-M, an *in silico* promoter analysis was carried out using the whole genomes of the three phytoplasmas ['*Ca*. P. asteris' strain AYWB (AYWB), '*Ca*. P. australianse' (PAa), and '*Ca*. P. mali' strain AT (ATP)]. Given that RpoD homologous proteins of these phytoplasmas contain highly conserved motifs, regions 2 and 4, which are responsible for the recognition of the –10 and –35 promoter elements, respectively (see Supplementary Fig. S2 online), their promoter recognition specificity would likely be similar to that of RpoD<sub>OY</sub>. Therefore, we used the same query sequences ([TC][AT][GC][AC][TC][AT] N<sub>17-19</sub>TA[AT]AA[AT]) for promoter prediction. In total, 68, 71, and 100 putative RpoD-dependent genes were found in AYWB, PAa, and ATP, respectively. Of the 88 putative RpoD<sub>OY</sub>-dependent genes in OY-M, 34 genes (39%) were also predicted to be RpoD-dependent in at least one of the other phytoplasma strains. Various housekeeping genes such as ribosomal RNA genes, ribosomal protein subunit genes, and tRNAs, were found in common among the putative RpoD-dependent genes of all three phytoplasmas (see Supplementary Table S3 online), implying that phytoplasmal RpoD plays a role, at least in part, as the principal sigma factor.

#### Discussion

All known eubacteria possess the principal sigma factor responsible for transcription of the majority of housekeeping genes. Hence, identification of the promoter elements recognized by the principal sigma factor is an important step toward understanding gene regulation mechanisms in bacteria. The principal sigma factors in culturable bacteria, such as *E. coli* RpoD and *Bacillus subtilis* SigA, have been well described, and their target promoter elements have been intensively studied genetically<sup>22,38,40</sup>. However, little is currently known about the sigma factors and their promoter sequences of obligate parasitic bacteria, including phytoplasmas, due to the difficulty of their *in vitro* culture and genetic engineering. In this study, we identified the RpoD-dependent promoters in OY-M phytoplasma by the development of an *in vitro* transcription assay system using an RNAP holoenzyme heterologously reconstituted with RNAP<sub>Ec</sub> and RpoD<sub>OY</sub>. This system would be also a powerful tool for studying gene regulatory mechanisms of other uncultured bacteria.

The in vitro transcription system has the advantage that transcripts synthesized from each promoter can be discriminated by their lengths. Based on the lengths of two major in vitro transcripts from PrrnB, which was the DNA template containing the putative *rrnB* promoter region (Fig. 1b), they were considered to be transcribed from promoter elements in the upstream region of P2 and P3 that were two out of the three rrnB TSSs determined by 5' RACE analysis (Fig. 1a). These putative promoter elements upstream of P2 (-35 5'-TTGCTA-3' and -10 5'-TATAAT-3') and P3 (-35 5'-TTGCCA-3' and -10 5'-TATAAT-3') were similar to the typical bacterial RpoD-dependent core promoter elements (-35 5'-TTGACA-3' and -10 5'-TATAAT-3'; Fig. 1a). Our results agree with earlier observations that in many bacteria, the upstream region of rrn has promoter elements that can be recognized by RpoD<sup>23,24</sup>. In particular, similar to rrnB of phytoplasmas, the rrn of E. coli and B. subtilis has two sets of promoter elements, which can also be recognized by sigma 70-type sigma factors<sup>21,22</sup>. However, the *rrnB* P1 promoter might be under the control of a sigma factor other than RpoD<sub>OV</sub> since no transcript corresponding to those from P1 was detected in the *in vitro* transcription system employing the  $RNAP_{Fc}$ -RpoD<sub>OV</sub> holoenzyme despite the fact that, in vivo, we identified the transcripts from P1 via the 5' RACE analysis. We also found nine genes (ibpA, infC, rplM, rpsD, PAM157, PAM289, PAM486, tengu, and amp) that have only one TSS with -10 and -35 promoter elements resembling the RpoD<sub>OV</sub>-dependent promoters of rrnB (P2 and P3) (Fig. 4a). Moreover, the predicted size from these identified TSSs to the 3'-ends of the *in vitro* transcription templates roughly corresponded to the length of transcripts in the RpoD<sub>OY</sub>-mediated in vitro transcription assays (Fig. 3). These results indicate that these genes are also most likely RpoD<sub>OY</sub>-dependent. If there were other sigma factors that transcribe these genes, two or more 5' end of *in vivo* transcripts would be identified.

Among them, three genes (rrnB, PAM289, and tengu) were common between the in vitro transcription assay in this study and another in vivo promoter assay<sup>19</sup>. The upstream regions of rrnB and PAM289 were recognized by  $RNAP_{Fc}$ -RpoD<sub>OY</sub> in our *in vitro* transcription assay (Fig. 3), which is consistent with the results of the EcERA system<sup>19</sup>. The *tengu* promoter recognized by  $RpoD_{OV}$  in our *in vitro* transcription (Fig. 3), however, exhibited no significant increases in transcription activity by RpoD<sub>OV</sub> in the EcERA system<sup>19</sup>. Our 5' RACE analysis using total RNA derived from phytoplasma-infected plants showed that the upstream region of major TSSs of *tengu* actually contains typical Rpo $D_{OV}$ -dependent promoter elements (-35 5'-TACATT-3' and -10 5'-TATAAT-3'; Fig. 4a), suggesting that tengu is also regulated in vivo by RpoD<sub>OY</sub>. Thus, the results of our in vitro transcription assay were not necessarily consistent with those obtained from the EcERA system. A possible explanation for this discrepancy between our results and those from the previous study is the side effects of several positive and negative transcriptional regulators of E. coli in the EcERA system. These regulators may affect the results of ECERA due to the interaction between these regulators of E. coli and target promoter sequences of phytoplasmas. Moreover, overexpression of RpoD<sub>OY</sub> can perturb the expression pattern of these regulators of E. coli because we revealed that RpoD<sub>OY</sub> recognizes promoter elements similar to those recognized by E. coli RpoD. Therefore, although promoters that require activation by other transcriptional regulators are likely to escape detection by our in vitro transcription system, it seems to be a more accurate tool compared to the EcERA system in measuring the specific activity of promoters recognized by RpoD<sub>OV</sub>.

Our *in vitro* transcription assay mediated by  $\text{RpoD}_{OY}$  revealed that nucleotide substitutions in the -10 promoter element of *rrnB* P2 drastically influenced promoter activity (Fig. 2d). In contrast, substitutions in the -35 promoter element, with the exception of the TT motif on the 5'-side, had little effect on promoter activity (Fig. 2b,c). Moreover, the -35 promoter elements of  $\text{RpoD}_{OY}$ -dependent phytoplasma genes were highly variable (Fig. 4b). In agreement with these findings, among the RpoD-dependent promoters of some other bacteria such as *E. coli, Campylobacter jejuni*, and *Mycoplasma hyopneumoniae*, the -10 promoter elements are very similar to each other, but the -35 promoter elements are relatively variable<sup>4,35,41</sup>. Nucleotide substitutions in the -35 promoter element of *Chlamydia trachomatis*, an obligate intracellular pathogen, had smaller effects compared to substitutions in the -10 promoter element of the RNAP holoenzyme<sup>24</sup>. In addition, in *M. hyopneumoniae*, which has a small AT-rich genome similar to phytoplasmas, no obvious -35 promoter elements were identified upstream of the TSSs of each gene, while the typical -10 promoter elements (5'-TATAAT-3') were found<sup>35</sup>. Thus, sequence features of the RpoD<sub>OY</sub>-dependent promoter elements identified in this study were consistent with previous observations of other bacteria.

In addition to the -35 and -10 promoter elements, the RpoD<sub>OY</sub>-dependent promoter possessed other sequence features common in bacteria, such as a 5'-TnTG-3' positioned around -17 to -14 ("the extended -10 region") and an A-rich region positioned around -42 to -39 (see Supplementary Fig. S5 online)<sup>36,37</sup>. In *E. coli*, the extended -10 region and the A-rich region were suggested to interact with region 3.0 (previously named 2.5) of RpoD and the  $\alpha$ -subunit of RNAP, respectively<sup>42,43</sup>. In the region 3.0, histidine and glutamic acid residues, which are conserved among bacterial sigma factors including phytoplasmal RpoDs (see Supplementary Fig. S2 online), are involved in contacting the extended -10 region<sup>42</sup>. Nucleotide substitution of the extended -10 region resulted in a drastic reduction of the promoter activity in *B. subtilis* and *E. coli*<sup>38,40</sup>. Moreover, in *C. trachomatis*, nucleotide substitutions at positions 4 and 5 bp upstream of TSSs, where no conserved motif has been identified, had negative effects on promoter activity<sup>24</sup>. Our *in vitro* transcription assay using RpoD<sub>OY</sub> revealed that nucleotide substitutions at positions -14 and -13 (mt11), and -6 and -5 (mt15), slightly decreased promoter activity (Fig. 2d). Examining the contribution of these extended regions of the phytoplasma promoter to transcriptional activity should be interesting.

In many bacteria, various sigma factors compete for a limited amount of RNAP core enzyme<sup>2</sup>, and the ratio of individual sigma factors could affect gene expression patterns. In general, the intracellular concentration of the principal sigma factor RpoD is held constant under ordinary conditions for transcription of the majority of the housekeeping genes, and other alternative sigma factors are transiently expressed under specific conditions for bacterial adaptation to environmental changes<sup>2,26</sup>. However, *rpo-* $D_{OY}$  expression is approximately four times more abundant in insect hosts compared to plant hosts in the phytoplasma life cycle<sup>11,19</sup>, while the expression level of *fliA*<sub>OY</sub>, another type of sigma factor gene in OY-M, does not differ significantly between the two host types<sup>19</sup>.

We had initially hypothesized that  $\text{RpoD}_{OY}$  would regulate the expression of genes that are highly expressed in insect hosts as the alternative sigma factor rather than the principal sigma factor<sup>11</sup>. Contrary to this hypothesis, our *in silico* search for  $\text{RpoD}_{OY}$ -dependent genes suggested that  $\text{RpoD}_{OY}$  mediated the transcription of many housekeeping genes as the principal sigma factor (see Supplementary Table S3 and Fig. S6 online). In addition, our *in vitro* transcription assays revealed that  $\text{RpoD}_{OY}$  recognized the promoters of *tengu* and *PAM486* (Fig. 3), which are highly expressed in plant hosts. In support of these results, the 5' RACE analysis using the total RNA of phytoplasma-infected plants showed that the upstream region of the major TSS of *tengu* and *PAM486* contained typical  $\text{RpoD}_{OY}$ -dependent promoter elements (Fig. 4). Therefore,  $\text{RpoD}_{OY}$  seems to mediate transcription of not only several housekeeping genes as the principal sigma factor, which is in agreement with classical theory, but also genes highly expressed in either plant or insect hosts.

How RpoD<sub>OY</sub>-dependent genes such as *tengu* and *PAM486* are highly expressed in plants is still uncertain, despite the fact that  $rpoD_{OY}$  is highly expressed in insects. We propose two hypotheses to explain these observations. First, other phytoplasma transcriptional regulators may positively or negatively affect the gene expression levels. For example, histone-like protein (HimA) conserved in phytoplasmas is one of the candidate transcription regulators. In some bacteria, histone-like proteins act as a transcriptional repressor by binding to DNA<sup>44,45</sup>. When phytoplasmas infect insect hosts, HimA may repress the expression of some of the RpoD<sub>OY</sub>-dependent genes, which leads to their specific expression in plant hosts. Second, non-coding RNAs may play roles as gene regulatory factors. Recent studies have revealed that non-coding RNAs such as riboswitches and small RNAs regulate their gene expression in many bacteria<sup>46,47</sup>. In a previous study, we used 5' RACE analysis to show frequent transcription initiation within the coding regions of genes<sup>48</sup>. We found in this study, using *in silico* promoter analysis, many promoter elements that positioned intergenic regions far from the ATG initiation codon or within the coding regions of genes, implying the existence of many non-coding RNAs in phytoplasmas.

In class *Mollicutes*, mycoplasmas and spiroplasmas possess *rpoD* as a single sigma factor gene in their genomes<sup>49,50</sup>. Like phytoplasmas, however, many mycoplasma genes are differentially expressed under various conditions<sup>51,52</sup>, and several spiroplasmas multiply in distinct plant and insect hosts<sup>53</sup>. Further studies of these bacteria will provide insight into the gene expression mechanisms of phytoplasmas.

Our *in silico* analysis detected 88 RpoD<sub>OY</sub>-dependent genes in the OY-M genome. Some of these genes, such as *znuA*, rRNA-16S, and tRNA-Glu, were known as the first genes of operons that could be regulated by the principal sigma factor in other bacteria<sup>21,22,54,55</sup>. In phytoplasmas, operon structures initiated from these genes are also conserved. Therefore, given the presence of these and other RpoD<sub>OY</sub>-dependent operons, RpoD<sub>OY</sub> is likely to regulate more than 88 genes identified by our *in silico* analysis, even though additional studies are needed to determine whether each mRNA of putative RpoD<sub>OY</sub>-dependent genes is monocistronic or polycistronic.

Numerous studies have implied that the principal sigma factor gene is transcribed by the RNAP holoenzyme containing the principal sigma factor itself <sup>56–58</sup>. However, in *Streptomyces griseus*, the principal sigma factor *hrdB* is not controlled by itself, but by the alternative sigma factor ShbA, which is known as an ECF sigma factor<sup>59</sup>. The ECF-subfamily sigma factors, the activity of which is regulated at the post-transcriptional level, usually activate the transcription of specific genes in response to environmental changes<sup>60</sup>. Although our results suggested a role of RpoD<sub>OY</sub> as the principal sigma factor, *PrpoD* was not controlled by RNAP<sub>Ec</sub>-RpoD<sub>OY</sub> (Fig. 3). The possibility exists that *rpoD<sub>OY</sub>* is transcribed as polycistronic mRNAs, but we could not find the upstream genes of *rpoD<sub>OY</sub>* (e.g., the PAM624 and glycyl-tRNA synthetase genes) that can constitute a set of operons containing *rpoD<sub>OY</sub>* from the list of putative RpoD<sub>OY</sub>-dependent genes (see supplementary Table S1 online). These findings imply that  $rpoD_{OY}$  transcription is regulated by a sigma factor other than RpoD<sub>OY</sub>, such as the putative ECF-like sigma factor FliA<sub>OY</sub>, raising the intriguing possibility that the sigma factor that transcribes  $rpoD_{OY}$  may play a key role as an environmental sensor because  $rpoD_{OY}$  expression is altered through host switching. Further studies are necessary to assess the effects of transcription activity of other transcriptional regulators such as FliA<sub>OY</sub> using *in vitro* transcription assays, which can reveal the detailed infection mechanisms of phytoplasma host switching between plants and insects.

### **Materials and Methods**

**Preparation of phytoplasma-infected plants and insects.** The '*Ca.* P. asteris' OY strain (OY) was isolated in Saga Prefecture, Japan<sup>61</sup>. A derivative line of OY (OY-M) was maintained in garland chrysan-themum (*Chrysanthemum coronarium*) using the leafhopper vector insect *Macrosteles striifrons*<sup>62</sup>. Plants infected with OY-M produce many lateral shoots, but exhibit only mild leaf yellowing and almost no stunting. OY-M-infected host plants exhibiting typical symptoms were maintained at 25 °C in a greenhouse with a 16-h light/8-h dark photoperiod until use for analysis. For total RNA extraction from insects, OY-M-carrying leafhoppers that fed on OY-M-infected plants for 40 days were used.

**5' RACE analysis and promoter prediction.** To identify the 5'-end of mRNA in OY-M, 5' RACE analysis was performed using the 5' RACE System for Rapid Amplification of cDNA Ends (Life Technologies Inc.) according to the manufacturer's protocol. Polymerase chain reaction (PCR) amplification was accomplished using Taq DNA polymerase (TaKaRa Bio Inc.), a nested gene-specific primer (RACE1 or RACE2) that anneals to a site located within the cDNA molecule (see Supplementary Table S1 online), and an anchor primer (Life Technologies Inc.). The nested PCR products were visualized by agarose gel electrophoresis and cloned using the pCR 2.1 TOPO® TA Cloning® Kit (Life Technologies, Inc.). Nucleotide sequences were determined by the dideoxynucleotide chain-termination method using an ABI Prism® 3130 Genetic Analyzer (Applied Biosystems). We sequenced at least eight independent clones and defined the most common 5'-end as a putative transcriptional start site of the genes. The consensus of the OY-M gene promoters was predicted using BioProspector<sup>63</sup>, available at http:// ai.stanford.edu/~xsliu/BioProspector/. The resulting consensus promoter sequence was illustrated based on multiple sequence alignments using the WebLogo tool<sup>64</sup>, available at http://weblogo.berkeley.edu/.

**Plasmid construction and expression of His-RpoD**<sub>OY</sub>. To construct pCold\_His-RpoD<sub>OY</sub>, the full-length *rpoD* gene of OY-M was PCR-amplified using a primer pair *rpoD*\_kpnF (5'-GGG GTA CCA TGG AAT TCG ATA ACA TAA TCA AAA -3') and *rpoD*\_salR (5'-CGA CGT CGA CTT ATT TGT GGT TGT GGT ACA AAC TTT TT-3'). The amplified DNA fragments were digested with *KpnI* and *SalI*, followed by cloning into pColdI (TaKaRa Bio Inc.) digested with the same enzymes. The resulting plasmid was subsequently transformed into *E. coli* strain BL21-CodonPlus<sup>TM</sup> (DE3)-RIL cells (Stratagene). The transformed *E. coli* was precultured at 37 °C in lysogeny broth (LB) medium containing ampicillin (50µg/ml). The overnight culture (2 ml) was added to 100 ml LB medium containing ampicillin (50µg/ml). After incubation for 1 h at 37 °C and subsequent incubation for 30 min at 15 °C, protein expression was induced by the addition of 0.1 mM isopropyl- $\beta$ -D-thiogalactopyranoside (IPTG). Cells were cultured at 15 °C for 24 h following induction.

In vitro transcription assay. N-terminally His-tagged-RpoD<sub>OY</sub> was purified as previously described<sup>65</sup> after overexpression in E. coli from pCold\_His-RpoD<sub>OY</sub>. A DNA-dependent RNA polymerase (RNAP) holoenzyme containing RpoD<sub>OY</sub> (RNAP<sub>Ec</sub>-RpoD<sub>OY</sub>) was made by adding 20 ng of purified RpoD<sub>OY</sub> to 1 U of E. coli RNAP core enzyme (Epicenter) in buffer [50 mM Tris-HCl, 100 mM KCl, 10 mM MgCl, 1 mM dithiothreitol, 0.1 mM EDTA, 5% (v/v) glycerol, pH 8.0] overnight at 4 °C. The reconstituted complex was verified by NativePAGE<sup>™</sup> gel electrophoresis (Life Technologies Inc.; see Supplementary Fig. S4 online). Transcription was initiated by adding 22 µl of NTP mix (500 µM ATP, 500 µM UTP, 500 µM GTP, 40 µM CTP; TaKaRa Bio Inc.), 2 µl of [alpha-32P]CTP (800 Ci/mmol; PerkinElmer), and 800 ng of DNA template to the reconstituted RNAP holoenzyme under the same buffer conditions. DNA fragments used for a template were amplified by PCR using the primer pair (see Supplementary Table S2 online) and purified using an UltraClean® 15 DNA Purification Kit (MO BIO Laboratories). A series of substitution templates were obtained via a recombinant PCR method<sup>66</sup>. Reaction mixtures were incubated for 15 min at 37 °C, and reactions were stopped by incubation for 5 min at 80 °C. Samples were treated with DNaseI for 15 min, followed by phenol-chlorophorm extraction and ethanol precipitation, and resuspended in 15µl loading dye containing 7.5M urea. Purified transcripts were analyzed using 6% (w/v) polyacrylamide-7 M urea gel electrophoresis and autoradiography. Signal intensities from autoradiographs were determined with the FLA-5000 image reader (GE Healthcare). All experiments were repeated at least twice and consistent results were obtained among replicates.

*In silico* prediction of RpoD<sub>OY</sub>-dependent genes in phytoplasma genomes. Promoter predictions were made in the genomes of *Ca*. P. asteris' OY-M, AYWB, *Ca*. P. australianse, and *Ca*. P. mali' AT (GenBank Accession Numbers AP006628.2, CP000061, AM422018, and CU469464, respectively) using genome sequence files obtained from the NCBI genome database (http://www.ncbi.nlm.nih.gov/genome/). Whole genome sequences were searched for promoter sequence patterns using 'dna-pattern'

of RSA-tools (http://rsat.ulb.ac.be/index.html). The resulting sequences were matched to each reference genome sequence, and the genes to which the resulting sequences were matched within 500 bp upstream of either ATG or 5'-end of mature rRNA and tRNA are listed in Supplementary Table S3 online.

#### References

- 1. Wösten, M. M. Eubacterial sigma-factors. FEMS Microbiol Rev 22, 127-150 (1998).
- Jishage, M., Iwata, A., Ueda, S. & Ishihama, A. Regulation of RNA polymerase sigma subunit synthesis in *Escherichia coli*: intracellular levels of four species of sigma subunit under various growth conditions. *J Bacteriol* 178, 5447–5451 (1996).
- 3. Gruber, T. M. & Gross, C. A. Multiple sigma subunits and the partitioning of bacterial transcription space. *Annu Rev Microbiol* **57**, 441–466 (2003).
- 4. Helmann, J. D. & Chamberlin, M. J. Structure and function of bacterial sigma factors. Annu Rev Biochem 57, 839-872 (1988).
- 5. Lee, I. M., Davis, R. E. & Gundersen-Rindal, D. E. Phytoplasma: phytopathogenic mollicutes. *Annu Rev Microbiol* 54, 221–255 (2000).
- 6. Maejima, K., Oshima, K. & Namba, S. Exploring the phytoplasmas, plant pathogenic bacteria. *J Gen Plant Pathol* **80**, 210–221 (2014).
- 7. Jackson J. A. Analysis of parasite host-switching: limitations on the use of phylogenies. Parasitology 199, S111–S123 (1999).
- 8. Hogenhout, S. A. et al. Phytoplasmas: bacteria that manipulate plants and insects. Mol Plant Pathol 9, 403-423 (2008).
- 9. Hoshi, A. et al. A unique virulence factor for proliferation and dwarfism in plants identified from a phytopathogenic bacterium. Proc Natl Acad Sci USA 106, 6416-6421 (2009).
- Maejima, K. et al. Recognition of floral homeotic MADS domain transcription factors by a phytoplasmal effector, phyllogen, induces phyllody. Plant J 78, 541–554 (2014).
- 11. Oshima, K. *et al.* Dramatic transcriptional changes in an intracellular parasite enable host switching between plant and insect. *PLoS One* **6**, e23242 (2011).
- 12. Oshima, K. *et al.* Reductive evolution suggested from the complete genome sequence of a plant-pathogenic phytoplasma. *Nat Genet* **36**, 27–39 (2004).
- 13. Bai, X. et al. Living with genome instability: the adaptation of phytoplasmas to diverse environments of their insect and plant hosts. J Bacteriol 188, 3682–3696 (2006).
- 14. Kube, M. et al. The linear chromosome of the plant-pathogenic mycoplasma 'Candidatus Phytoplasma mali'. BMC Genomics 9, 306 (2008).
- Tran-Nguyen, L. T., Kube, M., Schneider, B., Reinhardt, R. & Gibb, K. S. Comparative genome analysis of "Candidatus Phytoplasma australiense" (subgroup tuf-Australia I; rp-A) and "Ca. Phytoplasma asteris" Strains OY-M and AY-WB. J Bacteriol 190, 3979–3991 (2008).
- Andersen M. T., Liefting L. W., Havukkala I. & Beever R. E. Comparison of the complete genome sequence of two closely related isolates of '*Candidatus* Phytoplasma australiense' reveals genome plasticity. *BMC Genomics* 14, 529 (2013).
- 17. Hu, Y. & Coates, A. R. Transcription of two sigma 70 homologue genes, sigA and sigB, in stationary-phase Mycobacterium tuberculosis. J Bacteriol 181 469–476 (1999).
- 18. Ishihama, A. Functional modulation of Escherichia coli RNA polymerase. Annu Rev Microbiol 54, 499-518 (2000).
- 19. Ishii, Y., Kakizawa, S. & Oshima, K. New *ex vivo* reporter assay system reveals that  $\sigma$  factors of an unculturable pathogen control gene regulation involved in the host switching between insects and plants. *MicrobiologyOpen* **2**, 553–565 (2013).
- 20. Jung, H. Y. et al. First complete nucleotide sequence and heterologous gene organization of the two rRNA operons in the phytoplasma genome. DNA Cell Biol 22, 209–215 (2003).
- 21. Young, R. A. & Steitz, J. A. Tandem promoters direct E. coli ribosomal RNA synthesis. Cell 17, 225-234 (1979).
- 22. Ogasawara, N., Moriya, S. & Yoshikawa, H. Structure and organization of rRNA operons in the region of the replication origin of the *Bacillus subtilis* chromosome. *Nucleic Acids Res* **11**, 6301–6318 (1983).
- 23. Gourse, R. L., de Boer, H. A. & Nomura, M. DNA determinants of rRNA synthesis in *E. coli*: growth rate dependent regulation, feedback inhibition, upstream activation, antitermination. *Cell* 44, 197–205 (1986).
- Tan, M., Gaal, T., Gourse, R. L. & Engel, J. N. Mutational analysis of the *Chlamydia trachomatis* rRNA P1 promoter defines four regions important for transcription in vitro. J Bacteriol 180, 2359–2366 (1998).
- 25. Mantovani, R. An RNA polymerase II in vitro transcription system. Methods Mol Biol 31, 289-298 (1994).
- 26. Aramaki, H. & Fujita, M. In vitro transcription analysis of rpoD in Pseudomonas aeruginosa PAO1. FEMS Microbiol Lett 180, 311-316 (1999).
- Liu, J. & Zuber, P. The ClpX protein of *Bacillus subtilis* indirectly influences RNA polymerase holoenzyme composition and directly stimulates sigma-dependent transcription. *Mol Microbiol* 37, 885–897 (2000).
- Choonee, N., Even, S., Zig, L. & Putzer, H. Ribosomal protein L20 controls expression of the Bacillus subtilis infC operon via a transcription attenuation mechanism. Nucleic Acids Res 35, 1578–1588 (2007).
- Rosenthal, A. Z., Kim, Y. & Gralla, J. D. Regulation of transcription by acetate in *Escherichia coli: in vivo* and *in vitro* comparisons. *Mol Microbiol* 68, 907–917 (2008).
- 30. Stoll, S., Feldhaar, H. & Gross, R. Promoter characterization in the AT-rich genome of the obligate endosymbiont "Candidatus Blochmannia floridanus." J Bacteriol 191, 3747–3751 (2009).
- Wilcox, J. L., Dunbar, H. E., Wolfinger, R. D. & Moran, N. A. Consequences of reductive evolution for gene expression in an obligate endosymbiont. *Mol Microbiol* 48, 1491–1500 (2003).
- 32. Dunbar, H. E., Wilson, A. C., Ferguson, N. R. & Moran, N. A. Aphid thermal tolerance is governed by a point mutation in bacterial symbionts. *PLoS Biol* 5, e96 (2007).
- 33. Neriya, Y. *et al.* Onion yellow phytoplasma P38 protein plays a role in adhesion to the hosts. *FEMS microbiol lett* **361**, 115–122 (2014).
- 34. Suzuki, S. *et al.* Interaction between the membrane protein of a pathogen and insect microfilament complex determines insectvector specificity. *Proc Natl Acad Sci USA* **103**, 4252–4257 (2006).
- 35. Weber, Sde S., Sant'Anna, F. H. & Schrank, I. S. Unveiling *Mycoplasma hyopneumoniae* promoters: sequence definition and genomic distribution. *DNA Res* 19, 103–115 (2012).
- Graves, M. C. & Rabinowitz, J. C. In vivo and in vitro transcription of the Clostridium pasteurianum ferredoxin gene. Evidence for "extended" promoter elements in gram-positive organisms. J Biol Chem 261, 11409–11415 (1986).
- Helmann, J. D. Compilation and analysis of *Bacillus subtilis* sigma A-dependent promoter sequences: evidence for extended contact between RNA polymerase and upstream promoter DNA. *Nucleic Acids Res* 23, 2351–2360 (1995).
- Mitchell, J. E., Zheng, D., Busby, S. J. & Minchin, S. D. Identification and analysis of 'extended -10' promoters in *Escherichia coli*. *Nucleic Acids Res* 31, 4689–4695 (2003).
- Tatusov, R. L. et al. The COG database: new developments in phylogenetic classification of proteins from complete genomes. Nucleic Acids Res 29, 22–28 (2001).

- Voskuil, M. I., Voepel, K. & Chambliss, G. H. The -16 region, a vital sequence for the utilization of a promoter in Bacillus subtilis and Escherichia coli. Mol Microbiol 17, 271–279 (1995).
- Wösten, M. M., Boeve, M., Koot, M. G., van Nuenen, A. C. & van der Zeijst, B. A. Identification of *Campylobacter jejuni* promoter sequences. J Bacteriol 180, 594–599 (1998).
- 42. Barne, K. A., Bown, J. A., Busby, S. J. & Minchin, S. D. Region 2.5 of the *Escherichia coli* RNA polymerase sigma70 subunit is responsible for the recognition of the 'extended-10' motif at promoters. *EMBO J* 16, 4034–4040 (1997).
- Hook-Barnard, I. G. & Hinton, D. M. Transcription initiation by mix and match elements: flexibility for polymerase binding to bacterial promoters. *Gene Regul Syst Bio* 1, 275–293 (2007).
- 44. Atlung, T., Sund, S., Olesen, K. & Brøndsted, L. The histone-like protein H-NS acts as a transcriptional repressor for expression of the anaerobic and growth phase activator AppY of *Escherichia coli*. J Bacteriol **178**, 3418–3425 (1996).
- 45. Navarre, W. W. et al. Selective silencing of foreign DNA with low GC content by the H-NS protein in Salmonella. Science 313, 236–238 (2006).
- Sesto, N., Wurtzel, O., Archambaud, C., Sorek, R. & Cossart, P. The excludon: a new concept in bacterial antisense RNA-mediated gene regulation. *Nature reviews* 11, 75–82 (2013).
- 47. Lazarev, V. et al. Complete genome and proteome of Acholeplasma laidlawii. J bacteriol 193, 4943-4953 (2011).
- Ishii, Y. et al. In the non-insect-transmissible line of onion yellows phytoplasma (OY-NIM), the plasmid-encoded transmembrane protein ORF3 lacks the major promoter region. Microbiology 155, 2058–2067 (2009).
- 49. Razin, S., Yogev, D. & Naot, Y. Molecular biology and pathogenicity of mycoplasmas. *Microbiol Mol Biol Rev* 62, 1094–1156 (1998).
- 50. Ku, C., Lo, W. S., Chen, L. L. & Kuo, C. H. Complete genomes of two dipteran-associated spiroplasmas provided insights into the origin, dynamics, and impacts of viral invasion in spiroplasma. *Genome Biol Evol* **5**, 1151–1164 (2013).
- 51. Madsen, M. L., Puttamreddy, S., Thacker, E. L., Carruthers, M. D. & Minion, F. C. Transcriptome changes in *Mycoplasma hyopneumoniae* during infection. *Infect Immun* **76**, 658–663 (2008).
- 52. Zhang, W. & Baseman, J. B. Transcriptional response of *Mycoplasma genitalium* to osmotic stress. *Microbiology* 157, 548–556 (2011).
- 53. Gasparich, G. E. Spiroplasmas and phytoplasmas: microbes associated with plant hosts. Biologicals 38, 193-203 (2010).
- 54. Nilsson, D. & Johansen, E. A conserved sequence in tRNA and rRNA promoters of *Lactococcus lactis*. *Biochim Biophys Acta* **1219**, 141–144 (1994).
- Shin, J. H., Oh, S. Y., Kim, S. J. & Roe, J. H. The zinc-responsive regulator Zur controls a zinc uptake system and some ribosomal proteins in *Streptomyces coelicolor* A3(2). J Bacteriol 189, 4070–4077 (2007).
- Wang, L. F. & Doi, R. H. Promoter switching during development and the termination site of the sigma 43 operon of *Bacillus subtilis*. Mol Gen Genet 207, 114–119 (1987).
- 57. Fujita, M., Hanaura, Y. & Amemura, A. Analysis of the *rpoD* gene encoding the principal sigma factor of *Pseudomonas putida*. *Gene* **167**, 93–98 (1995).
- Rodrigue, S. et al. Identification of mycobacterial sigma factor binding sites by chromatin immunoprecipitation assays. J Bacteriol 189, 1505–1513 (2007).
- Otani, H., Higo, A., Nanamiya, H., Horinouchi, S. & Ohnishi, Y. An alternative sigma factor governs the principal sigma factor in *Streptomyces griseus*. Mol Microbiol 87, 1223–1236 (2013).
- 60. Mascher, T. Signaling diversity and evolution of extracytoplasmic function (ECF)  $\sigma$  factors. Curr Opin Microbiol 16, 148–155 (2013).
- 61. Shiomi T., Tanaka M., Waki H. & Zenbayashi R. Occurrence of welsh onion yellows. Ann Phytopathol Soc Jpn 62, 258-260 (1996).
- Oshima, K. et al. Isolation and characterization of derivative lines of the onion yellows phytoplasma that do not cause stunting or phloem hyperplasia. *Phytopathology* 91, 1024–1029 (2001).
- Liu, X., Brutlag, D. L. & Liu, J. S. BioProspector: Discovering conserved DNA motifs in upstream regulatory regions of co-expressed genes. *Pac Symp Biocomput* 6, 127–138 (2001).
- 64. Crooks, G. E., Hon, G., Chandonia, J. M. & Brenner, S. E. WebLogo: a sequence logo generator. *Genome Res* 14, 1188–1190 (2004). 65. Kakizawa, S. *et al.* Cloning of immunodominant membrane protein genes of phytoplasmas and their in planta expression. *FEMS*
- microbiol lett **293**, 91–101 (2009). 66. Higuchi, R. in PCR Protocols: A Guide to Methods and Applications. (eds Innis, M. A. *et al.*) Ch. 22, 177–183 (Academic Press,
- Higuchi, R. in PCR Protocols: A Guide to Methods and Applications. (eds Innis, M. A. et al.) Ch. 22, 177–183 (Academic Press, 1990).

#### Acknowledgments

This work was supported in part by the Japan Society for the Promotion of Science (JSPS) through grantsin-aids for JSPS fellows and for Scientific Research, the Funding program for Next Generation World-Leading Researchers (project GS005) initiated by the Council for Science and Technology Policy (CSTP), and the Program for the Promotion of Basic Research Activities for Innovative Bioscience (PROBRAIN).

#### **Author Contributions**

C.M., K.K., K.M., M.H., K.O. and S.N. designed research; C.M., K.K., T.N., Y.K., T.T. and A.Y. performed research; and C.M. wrote the main manuscript text. All authors discussed the results and commented on the manuscript.

### **Additional Information**

Supplementary information accompanies this paper at http://www.nature.com/srep

Competing financial interests: The authors declare no competing financial interests.

How to cite this article: Miura, C. *et al.* Functional characterization of the principal sigma factor RpoD of phytoplasmas *via* an *in vitro* transcription assay. *Sci. Rep.* **5**, 11893; doi: 10.1038/srep11893 (2015).

This work is licensed under a Creative Commons Attribution 4.0 International License. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in the credit line; if the material is not included under the Creative Commons license, users will need to obtain permission from the license holder to reproduce the material. To view a copy of this license, visit http://creativecommons.org/licenses/by/4.0/