Diversity of the HrpF/NolX/PopF protein in plant pathogenic bacteria

E.I. Kyrova

People's Friendship University of Russia, Miklukho-Maklaya 6, Moscow, Russia, e-mail:esheverdina@yandex.ru

J. Panycheva

"PhytoEngineering" R&D Center, LLC, Moscow region, Russia, e-mail: yulya_panycheva@mail.ru

A.N. Ignatov

"PhytoEngineering" R&D Center, LLC, Moscow region; M.M. Shemyakin–Yu.A. Ovchinnikov Institute of Bioorganic Chemistry of the Russian Academy of Sciences, Moscow, Russia, e-mail:an.ignatov@gmail.com

Introduction The interaction between the Gramm-negative plant pathogenic bacteria and its host plants is controlled by *hrp* genes (*hypersensitive reaction and pathogenicity*), which encode a type III protein secretion system (T3SS). T3SS-secreted proteins include hrp-related translocators and effectors involved in plant defense suppression. The HrpF/NolX/ HrpK/ PopF1/ PopF2 family of hrp proteins (PF05819) includes *Rhizobium* NolX, *Ralstonia* Pop1/2, and *Xanthomonas* HrpF proteins [1]. Despite large taxonomic distance between genera *Rhizobium* (alfa-proteobacteria), *Ralstonia* (beta-proteobacteria), and *Xanthomonas* (Gamma-proteobacteria), HrpF, NolX and PopF1/ PopF2 have remarkably high similarity [3,4].

HrpF is a translocator of effector proteins into the host cell, dispensable for protein secretion but required for effector AvrBs3 recognition in planta [2]. NolX, a legumes cultivar-specific protein, is expressed in planta mostly during the early stages of nodule development [3], secreted by a T3SS [4]. Expression of nolX is induced as much as 30 fold by flavonoid signal molecules, even though these genes lack nodbox promoters [5]. It was speculated that HrpF protein may play a role in host-specialization of the bacteria [6], and *hrpF* gene -specific primers were successfully used for identification of *X. campestris* [7].

We evaluated protein diversity across different HrpF/NolX/PopF1/PopF2 group members available in the GeneBank, and sequenced nearly complete *hrpF* gene fragment for 65 strains of xanthomonads obtained from brassicas with symptoms of black rot or leaf spots across Russia and some related bacteria isolated from crops of other families at the same regions. 91 proteins of HrpF/NolX/HrpK/PopF1/PopF2 family were found in the GeneBank by BLAST search against HrpF protein of *Xanthomonas campestris* pv. *campestris* str. ATCC 3391372. 65 phenotypically and genetically characterized strains of Xanthomonas from brassicas collected in Russia, and

representative strains of related bacteria from USA, Japan, Germany, and other countries were obtained from collection of Russian Research Institute of Phytopathology. Majority of the strains belonged to *X. campestris pv. campestris, pv. raphani, pv. armoraciae,* and included some strains of *X. arboricola, X. vesicatoria,* and *X. hortorum.* Nearly complete DNA sequence of *hrpF* gene (85%) was obtained with two pairs of sequencing primers by direct PCR amplification [8].

Results The HrpF proteins obtained for *Xanthomonas* spp., NolX, and PopF1/ PopF2 had a length ranged from 806 to 921aa. The protein is hydrophilic, but it has two hydrophobic domains at the C-terminal end. The protein sequence does not contain signal peptides. Molecular weight is ranged from 87 to 98.9 kDa. A detailed analysis of the obtained amino acid sequence revealed two repetitive regions in HrpF/NolX/Pop1 proteins. One of these is a conserved domain with 118aa length. In most proteins of the HrpF/Pop1 group this conserved domain is present in two copies, but in the NolX proteins it is single-copy, and in the HrpF protein from *X. campestris* the 188aa domain is represented by three copies and demonstrates a possible origin from recombination with the homologous protein from *X. fuscans*.

The second repetitive region is a 4aa degenerate repeat (consensus sequence AKGA). The first 4 repeats of amino acid sequence **AKGA** are coded by conservative DNA sequence 5'-CTC GCA GCC AAG GGA GCA GCC AAG GGG GTC GCC GAG GGC GCA GCC AAG GGA GCG ACC CAG GGT-3' present in many sequenced bacterial genomes. Some regulatory proteins can bind these sequences. We also assume that these repeats are responsible for the rapid evolution of this tandem repetitive region of the HrpF protein and may determine the range of host-plants.

In the HrpF protein of *X. campestris* it starts at 810aa position, and present from 8 to 28 times in tandem repeats in the homologous protein sequences from other bacteria of the genus *Xanthomonas*. We have noted the similarity of the four amino acid repeat with the EKGI/AKGI repeat of transposase, which is probably responsible for the DNA binding. The analysis of the obtained *hrpF* gene nucleotide sequences from xanthomonads isolated from brassicas mainly reflected the host plant range of the isolated bacteria then their taxonomic position obtained by MLST analysis (data not shown).

Conclusions HrpF/NolX/Pop1 proteins are probable determinants of host range of the plant pathogenic and symbiotic bacteria with nucleotide binding activity and specific mechanism of their rapid evolution.

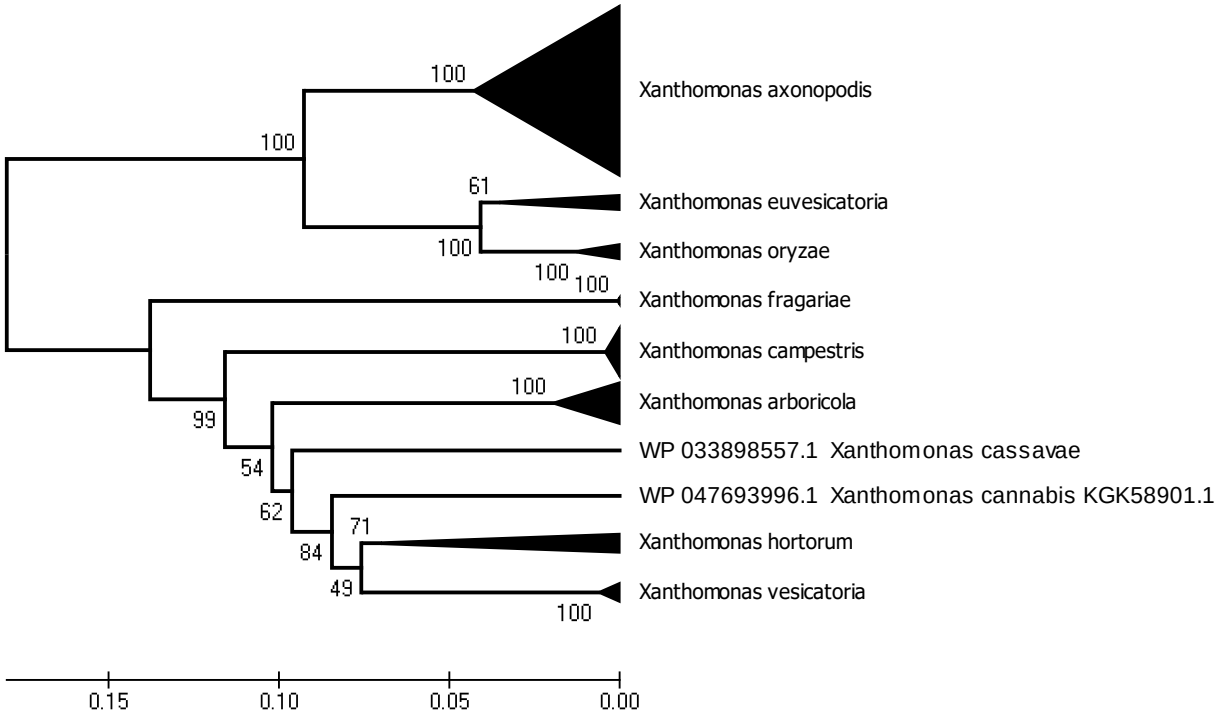

Figure 1. Evolutionary relationships of 91 proteins of HrpX/NolX/Pop1 group from *Xanthomonas* spp.. The evolutionary history was inferred using the Neighbor-Joining method. The percentage of replicate trees in which the associated taxa clustered together in the bootstrap test (1000 replicates) are shown next to the branches. Phylogenetic analyses were conducted in MEGA7.

All (110 aa) 1 (NolX) 2 (HrpF/Pop1) 3 (HrpF *X.campestris*) All (421 aa)

A B C D E

a)

>Consensus repeated sequence of HrpF protein:

F*PK*IKG*PPA**GS*VTW**GTLT*SEL*IVSTLN*HKD*********LD*KINDPSTPPD LK*AL*GL**DPRLFFAIGSQGDGKCGGK*****L*DF***H*Q******K*

b)

Figure 2. Conservative and variable repeated domens (B,C and D) of HrpF/NolX/Pop1 protein (a), and consensus repeated sequence of HrpF protein (b).

Acknowledgments This work was supported by research grant No 14.W03.31.0003 of Russian Government.

References

1. D. Meyer et al. (2006) PopF1 and PopF2, two proteins secreted by the type III protein secretion system of *Ralstonia solanacearum*, are translocators belonging to the HrpF/NopX family. *Journal of bacteriology*, ***188*****(13):** 4903-4917.
2. O. Rossier et al. (2000) HrpB2 and HrpF from *Xanthomonas* are type III-secreted proteins and essential for pathogenicity and recognition by the host plant. *Molecular microbiology*, ***38*****(4):** 828-838.
3. H.B. Krishnan (2002) NolX of *Sinorhizobium fredii* USDA257, a type III-secreted protein involved in host range determination, is localized in the infection threads of cowpea (Vigna unguiculata [L.] Walp) and soybean (*Glycine max* [L.] Merr.) nodules. *Journal of bacteriology*, **184(3):** 831-839.
4. E. Huguet and U. Bonas (1997) hrpF of Xanthomonas campestris pv. vesicatoria encodes an 87-kDa protein with homology to NolX of *Rhizobium fredii*. *Molecular plant-microbe interactions*, ***10*****(4):** 488-498.
5. L.W. Meinhardt et al. (1993) Molecular cloning and characterization of a sym plasmid locus that regulates cultivar-specific nodulation of soybean by *Rhizobium fredii* USDA257. *Molecular microbiology*, ***9*****(1):** 17-29.
6. F.J. Lopez-Baena et al. ero (2008) Regulation and symbiotic significance of nodulation outer proteins secretion in *Sinorhizobium fredii* HH103. *Microbiology*, ***154*****(6):** 1825-1836.
7. Y.J. Park et al. (2004) Sensitive and specific detection of Xanthomonas campestris pv. campestris by PCR using species-specific primers based on *hrpF* gene sequences. *Microbiological Research*, **159(4):**419-423.
8. A. Sugio et al. (2005). Characterization of the *hrpF* pathogenicity peninsula of Xanthomonas oryzae pv. oryzae. *Molecular plant-microbe interactions*, **18(6)**:546-554.