

Additional file 12: Table S5. Amino acid sequences of TAL effector proteins in *X. translucens* pv. *undulosa* strain XT4699.

XT4699-Tal1
MDPIRSRTFSPARELQAGSQPDGVQPTADPRVSPPAGSPLDGLPARRTMSRTQLPPFPASGPSFSAGSFSDLLRQVDSSSLFDAS FFDSMPAFGAHHAQAATGELDEAQSALRAVDDPQPSASAAITAAPRRTKAAARRCSAQTLDALPAADVLDSTFGYSQQQKEKIK PKVRSTVAQHHAALVGHGFTHAHIVELSKHPPALGTIAARYSEMIAALPEATHEDIVGVGKHCAGARTLEVLLMVVQELRAPPL QLVTSQLLKIAKRGGVTAVEAVHASRNALTGAPLH LLPDQVVAIVSHDGGKQSLETVQRLLPVLRQPPYG LTPNQVVAIASYDGGKQSLETVERLLPVLCQPPYG LTPNQVVAIASNIGGKPALETVQRLLPVLCQPPYG LTPNQVVAIASNNGGKQALETVQRLLPVLCQEYG LIPEQVVAIASNNGGKQALETVQRLLPVLCKEYG LTPEQVVAIASNNGGKLALETVERLLPVLCQPPYG LTPKQVVAIASYKGANQALGTVQRLLPVLCQPPYG LTPDQVVAIASNNGAKQALETVQRLLPVLCQPPYG LTPDQVVAIASHDGGRQALETVQRLLPVLCQPPYG LTPNQVVAIASNNGGKQALETVQRLLPVLCQEYG LTPNQVVAIASNNGGKQSLETVQRLLPVLCQPPYN LTPDQVVAIASNDGGKQALETVQRLLPVLCQPPYN LTPEQVVAIASNNGGKQALETVQRLLPVLCKEYG LTPEQVVAIASQDGGKQSLETVQRLLPVLCQPPYG LTPNQVVAIASNHGGKQSLETVQRLLPVLCQPPYG LTPNQVVAIASHDGGTQALE SIFAQLSSPDPALAALTNDRLVALACIGGRPALDAVKKGLPHAPALITRVHNRVPEGTAHLVADLAQVVRVLSFFQCHSHPAQA FDEAMRQFGMSRHLLQLFRRVGVTELEAISGLTPPASQRWDRMLQASGRKGAKPPSASAQTSQGESLDAFADSLERELDAPSP MHQAGQTLASSRKRSRSESSVNRSSAQQAEEVFVEQRDAPPLPLSSWGVKRRRTRIGGLPDPGTPTHGDLAASSAAFLEQDA DPFAGAAEDFPADFQEEIAWLKELLAH
XT4699-Tal2
MDPIRSRTFSPARELQAGSQPDGVQPTADPRVSPPAGSPLDGLPARRTMSRTQLPPFPASVPFASAGSFSDLLRQVDSSSLFDAS FFDSMPAFGAHHAQAATGELDEVQSALRAADDPQPPVRVAVTAARPPRAKPAQRPRRAAQTSASPAADVLDSTFGYSQQQKEK IKPTVRSSVAQHHAALVGHGFTHAHIVELSKHPAALGTIAARYSEMIAALPEATHEDIVGVGKQWSGARALEALLMVAEELRAP PLQLVTGQLLKIARKGGVTAVEAVHASRNALTGAPLH LTPDQVVAIVSNNGGKQALETVQRLLPVLCQPPYG LTPEQVVAIASHDGAKPALETVQRLLPVLCQPPYG LTPEQVVAIASNNGGKQALETVQRLLPVLCKEYG LTPEQVVAIASNNGGKLALETVERLLPVLCQPPYG LTPKQVVAIASHNGGKQALETVQRLLPVLCQPPYG LTPEQVVAIASKGGKQALETVQRLLPVLCQPPYG LTPDQVVTIANNIGAKQALETVQRLLPVLCQPPYG LTPNQVVAIASHDGAKQALETVQRLLPVLCQEYG LTPGQVVAIANNIGGKPALETVQRLLPVLCQPPYN LTPNQVVAIASNNGGKQALETVQRLLPVLCQPPHP LTPNQVVAIASHDGAKQALETVQRLLPVLCQSPYG LTPDQVVVIASHNGGKQALETVQRLLPVLCQPPYG LTPNQVVAIASHDGGKPALETVQRLLPVLCQPPYG LTPDQVVAIASHDGGKQALETVQRLLPVLCQDHG LTPGQVVAIANNIGGKPALETVQRLLPVLCKEYG LTPNQVVAIASHNGGKQALETVQRLLPVLCQPPYG LTPEQVVAIASHDGGKPALETVQRLLPVLCQPPYG LTPNQVVAIASQDGGTQALE SIFAQLSSPDPALAALTNDRLVALACIGGRPALDAVKKGLPHAPALITRVHNRVPEGTAHLVADHAQVVRVLSFFQCHSQRGQV FHEAMKRFEMSREGLLQLFRRVGVTELEAISGLTPPASQRWDRMLQASGRKGAKPPSASAQTSQGESLDAFADSLERELDAPSP MHQAGQTLASNRKRSRSESSVNRSSAQQAEEVFVEQRDAPPLPLSSWGVKRRRTRIGGLPDPGTPTHGDLAASSAAFLEQDA DPFAGAAEDFPADFQEEIAWLKELLAH

XT4699-Tal3

MDPIRSRTPSPARELQAGSQPDAVQPIADRLVSPPAGSPLDGLPARRTMSRTQLPSFPASVPAFSAGSFSDLLRQVDSSSLFDAS
FFDSMPAFGAHHAQAATGELDEVQSALRAADDPQPPVRVAVTAARPRAKPAQRPRRAAQTSDASPAADVLDSTFGYSQQQKEK
IKPTVVRSTVSQHHAALVGHGFTHAHIVELSKHPAALGTIAARYSEMIAALPEAAHEDIVGVGKQWSGARALEALLMVAEELRAP
PLQLVTGQLLKIAGGGVTAVEAVHASRNALTGAPLH
LTPDQVVVAIVSNNGGKQALETVQRLLPVLCQPPYN
LTPEQVVVAIASHDGAKPALETVQRLLPVLCQPPYG
LTPEQVVVAIASNNGGKQALEAVQRLLPVLCKEYG
LTPEQVVVAIASHDGAKQALETVQRLLLVLCQEYG
LTPNQVVVAIASHDGAKQALETVQRLLPVLCQPPYG
LTPNQVVVAIASHNGGKQALETVQRLLPVLCQPPYG
LTPEQVVVTIANNFEGGKPALETVQRLLPVLCQPPYG
LTPNQVVVAIANNIGAKPALETVQRLLPVLCQPPYG
LTPEQVVVAIAGNHGAKQALETVQRLLPVLCQPPYG
LTPNQVVVAIASHDGAKQALETVQRLLPVLCCKPPHP
LTPNQVVVAIASHDGAKQALETVQRLLPVLCCKPPYG
LTPDQVVVAIASHDGAKPALETVQRLLPVLCCKPPYG
LTPDQVVVAIASHNGGKQALETVQRLLPVLCQPSYG
LTRNQVVVAIASHNGGKQALETVQRLLPVLCKEYG
LTPEQVVVAIASHDGGTQALE
SIFAQLSSPDPALAALTNDRLVALACIGGRPALDAVKKGLPHAPALITRVHNRVPEGTAHLVADLAQVVRVLSFFQCHSHPAQA
FDEAMRQFGMSRHGLLQLFRRVGTELEAISGTLPPASQRWDRMLQASGRKGAKPPSASAQTTQGQESLDAFADSLERELDAPSP
MHQAGQTLASSRKRSRSESSVNRSSAQQAEEVFVPEQRDAPPLPLSSWGKRRRTRIGGLPDPGTPTHGDLAASSAAFLEQDA
DPFAGAAEDFPADFQEEIAWLKELLAH

XT4699-Tal4

MEPIRSRTPSTARELQAGSQPDAVQPIADRLVSTPASSPLDGLPARRMVSRTSPPSPARSPAFSAGSLSGLLRQQIDPSLFAG
SPFDSLPSFGAARASAPGEGDEVQSGLRVDDPQPSASAAITAAPPRTKAAARRRSAQTSDALPAAHVDLGTFGYSQQQKEKI
KPKVRSTLAQHHEALVGHGFTHAHIVELSKHPPALGTIAARYSEMIAALPEATHEDIVGVGKQKSGARALEALLTVAAEELRAPP
LQLVTGQLLKIARKGGVNAVEAVHASRNALTGAPLH
LLPDQVVVAIVSNHGSKLALGTVQRLLPVLCCKPPYG
LTRNQVVVAIVNNNGGKQALETVHRLLPVLCQPPYG
LTPEQVVVAIASHDGGRQALETVHRLLPVLRQPPYG
LTLEQVVVAIASNNGGKQALETVQRLLPVLCQPPYG
LTPDQVVVTIASHDGGRQALETVQRLLPVLCQPPYG
LTPNQVVVAIASNHGGKQALETVQRLLPVLCQPLHG
LTPDQVVVTIASHDGGRQALETVQRLLPVLCQPPYG
LTPKQVVVAIANYKGAKQALETVQRLLPLLCKPPYG
LTPDQVIAIVSNNGGKPALETVRRLPVLCCKHPYG
LTPKQVVVAIVSNHGGKPALETVRRLPVLCCKHPYG
LTPKQVVVAIASY*GGKPALETVQRLLPVLCQPPYG
LTPNQVVVAIASHDGGRQALETVQRLLPVLRQPPYG
LTPDQVVVAIASNNGGKQALETVQRLLPVLCQPPYS
LIPDQVVVAIASNIGAKQALETVQRLLPVLCQPPYG
LTTDQVIAIASNNGGSKQALETVRLLPVLCQPPYG
LTPDQVVVAIASQDGGKQALE
SIFAQLSSPDPALAALTNDRLVALACLGRPALDAVKKGLPHAPALVTRTHNRVPEGTAHLVADHAQVVRVLGFFQCHSHPAQA
FHEAMTRFEMSREGLLQLFRRVGTELEAISGTLPPASQRWHRIQLALGVKGAKPPSSASAQTPSQESVDAFADSLERELDAPSP
IHEADRARASNKRSRSESSVNRSSAQQAEEVFVPEQRDAPPLPLSSWGKVRQRTRIGGLPDPGMPTDGELAASSAAFLEQDA
DPFAGAAEDFPVFDQEEIAWLMTLLPH

XT4699-Tal5

MDPIRSRTSPSPARELQAGSQPDAVQPIADRLVSPAGSPLDGLPARRTMSRTQLPSPPASVPAFSAGSFSDLLRQVDSSSLFDAS
FFDSMPAFGAHHAQAATGELDEVQSALRAADDPQPPVRVAVTAARPPRAKPAQRPRRAAQTSDASPAADVLDSTFGYSQQQKEK
IKPTVRSTVAQHHAALVGHGFTHAHIVELSKHPPALGTIAARYSEMIAALPEATHEDIVGVGKQWSGARALDALLMVAEELRAP
PLQLVTGQLLKIAKRGVTAVEAVHASRNALTGAPLH
LTPDQVVAIVSNNGGKQALETVRLLPVLCQPPYG
LTPEQVVAIASHDGAKPALETVQRLLPVLCQPPYG
LTPEQVVAIASNNGGKPALETVQRLLPVLCKEYG
LTPEQVVAIASNNGGKLALETVERLLPVLCQPPYG
LTPNQVVAIASHNNGGKSLETVQRLLPVLCQPPYG
LTPEQVVAIASHNNGGKQALETVQRLLPVLCQPPYG
LTPEQVVTIANNIGGKQALETVQRLLPVLCQPPYG
LTPNQVVAIASNIGAKPALETVQRLLPVLCQPPYG
LTPEQVVTIANNIGGKPALETVQRLLPVLRKPPYG
LTPEQVVAIASNHGGKQALETVQRLLPVLCCKPPYG
LTPNQVVAIASNNGAKQALETVQRLLPVLCCKPPHP
LTPNQVVAIASHDGAKQALETVQRLLPVLCQPPYG
LTPEQVVVIASNNGGKQALETVQRLLPVLCQPSYG
LTPEQVVAIASNHGGKQALETVQRLLPVLCQPPYG
LTPNQVVAIASHDGKPALETVQRLLPVLCQPPYG
LTPNQVVAIASHDGGTQALE
SIFAQLSSPDPALAALTNDRLVALACIGGRPALDAVKKGLPHAPALITRVHNRVPEGTAHLVADLAQVVRVLSFFQCHSHPAQA
FDEAMRQFGMSRHGLLQLFRRVGVTLEAISGTLPPASQRWDRMLQASGRKGAKPSSASAQTPSQESVDAFADSLERELDAPSP
MHQAGQTLASSRKRSSESSVNRSSAQQAEEVFVPEQRDAPLLPLSSWGVRRRRTRIGGLPDPGTPTHGDLAASSGAFLEQDA
DPFAGAAEDFPADFQEEIAWLKELLAH

XT4699-Tal6

MDPIRSRTSPSPARELQAGSQPDAVQPIADRLVSPAGSPLDGLPARRTMSRTQLPSPPASVPAFSAGSFSDLLRQVDSSSLFDAS
FFDSMPAFGAHHAQAATGELDEVQSALRAADDPQPPVRVAVTAARPPRAKPAQRPRRAAQTSDASPAADVLDSTFGYSQQQKEK
IKPTVRSTVAQHHAALVGHGFTHAHIVELSKHPPAALGTIAARYSEMIAALPEATHEDIVGVGKQWSGARALETLLMVAEELRAP
PLQLVTGQLLKIAGGGVTAVEAVHASRNALTGAPLH
LTPDQVVAIVSHDGGKSLETVQRLLPVLCQPPYG
LTPNQVVAIASHNNGGKQALETVQRLLPVLCQEYG
LTPEQVVAIASHNNGGKQALETVQRLLPVLCQEYG
LTPEQVVAIASHDGAKQALETVQRLLPVLCQPPYG
LTPEQVVAIASNHGGKLALETVERLLPVLCQPPYG
LTPNQVVAIASNHGGKQALETVQPLLVLVCQPPYG
LTPEQVVAIASHGGAKQALKTVQRLLPVLCQDHG
LTPEQVVAIANHDGAKQALETVQRLLPVLCQPPYG
LTPEQVVAIASKGGGKQALETVQRLLPVLCQPPYG
LTPDQVVTIASNNGGKPALETVRLLPVLCCKPPYG
LTPKQVVAIASY*GGKQALETVQRLLPVLCCKPPYG
LTPEQVVAIASNNGGKPALETVQRLLPVLCQEYG
LTPNQVVAIASHDGAKPALETVQRLLPVLCQEYG
LTPNQVVAIASHDGKPALETVQRLLPVLCQPPYG
LTPEQVVAIASHNGGTQALE
SIFAQLSSPDPALAALTNDRLVALACIGGRPALDAVKKGLPHAPALITRVHNRVPEGTAHLVADLAQVVRVLSFFQCHSHPAQA
FDEAMRQFGMSRHGLLQLFRRVGVTLEAISGTLPPASQRWDRMLQASGRKAAKPPSASAQTPSQQESLDAFADSLERELDAPSP
MHQAGQTLASSRKRSSESSVNRSSAQQAEEVFVPEQRDAPLLPLSSWGVRRRRTRIGGLPDPGTPTHGDLAASSAAFLEQDA
DPFAGAAEDFPADFQEEIAWLKELLAH

XT4699-Tal7 MDPIRSRTSPSPARELLAGSQPDGVQPTADPRVSPPAGSPLDGLPARRTMSRTQLPPPSASGPAFSAGSFSDLLRQVDSSLFNAS LFDSMPAFGAHHAQAATGELDEAQSALRAVDDPQSSASAAITAAPRPTKAAARRRSAQTLSDASPAADVLDSTFGYSQQQQEKIK PTVRSSVAQHHAALVGHGFTHAHIVELSKHPAALGTIAARYSEMIAALPEATHEDIVGVGKQWSGARALEALLMVAEELRAPPL QLVTGQLLKIAKRGGVTAVEAVHASRNALTGAPLH LTPDQVVAIVSHNGGKQALETVQRLLPVLCQPPYN LTPEQVVAIASHDGGKQALETVQRLLLVLCQECG LTPNQVVAIASHDGAKPALETVHRLLPVLCQPPYG LTPNQVVAIASHDGAKQALETVQRLLPVLCQDHG LTPGQVVAIADNIGGKQALETVQRLLPVLCCKPPYG LTPEQVVTIANNIGGKPALETVQRLLPVLCCKPPYG LTPEQVVTIANNIGAKPALETVHRLLPVLRKPPYG LTPEQVVAIASHNGAKPALETVQRLLPVLCQPPYG LTPNQVVAIASHDGAKQALETVQRLLPVLCCKPPHP LTPNQVVAIASHDGAKQALETVQRLLPVLCQPPYG LTPDQVVVIASNNGGKQALETVQRLLPVLCQPPYG LTPEQVVAIASNNGGKPALETVQRLLPVLCQPPYD LTPDQVVAIANNIGAKPALETVQRLLPVLCQPPYG LTPEQVVAIASNNGGKQALETVQRLLPVLCQPPYG LTPYQVVAIASHDGGTQALE SIFAQLSSPDPALATLTNDRLVALACIGGRPALDAVKKGLPHAPELITRVHNRVPEGTAHLVADHAQVVRVLGFFQCHSQRGQV FHEAMKRFEMSREGLLQLFRRVGTELEAISGTLPPASQRWDRMLQASGRKGAKPPSASAQTQGQESLDAFADSLERELDAPSP MHQAGQTLASSRKRSRSESSVNRSSAQQAEEVFVPEQRDAPLLPLSSWGVRRRRTRIGGLPDTGMPTDGELAASSAAFLEQDA DPFAGAAEDFPADFQEEIAWLRELLAH
XT4699-Tal8 MDPIRSRTSPSPARELQAGSQPDVQPIADRLVSPPAGSPLDGLPARRTMSRTQLPSPPASVPAFSAGSFSDLLRQVDSSLFDAS FFDSMPAFGAHHAQAATGELDEVQSALRAADDPQPPVRVAVTAARPPRAKPAQRPRRAAQTSDASPAADVLDSTFGYSQQQQEK IKPTVRSTVAQHHAALVGHGFTHAHIVELSKHPAALGTIAARYSEMIAALPEATHEDIVGVGKQWSGARALEALLMVAEELRAP PLQLVTGQLLKIAKRGGVTAVEAVHASRNALTGAPLH LTPDQVVAIVSNNGGKQALETVQRLLPVLCCKPPYG LTPEQVVAIASNNGGKPALETVQRLLPVLCQEYG LTPNQVVAIASHDGAKQALETVQRLLPVLCQPPHP LTPNQVVAIASHDGGKPALETVQRLLPVLCCKPPYG LTPEQVVAIANHDGAKQALETVQRLLPVLCQPPYG LTPEQVVAIASKGGGKQALETVQRLLPVLCQPPYG LTPDQVVTIASNNGGKPALETVRRLLPVLCCKPPYG LTPKQVVAIASY*GGKQSLETVQRLLPVLCCKPPYG LTPEQVVAIASNNGGKPALETVQRLLPVLCQEYG LTPNQVVAIASHDGAKPALETVQRLLPVLCQEYG LTPNQVVAIASHDGAKQALETVQRLLPVLCCKPPYG LTPEQVVAIASQDGGKQSLETVQRLLPVLCCKPPYG LTPEQVVAIASHNGGTQALE SIFAQLSSPDPALAALTNGRLVALACIGGRPALDAVKKGLPHAPALITRVHNRVPEGTAHLVADLAQVVRVLSFFQCHSHPAQA FDEAMRQFGMSRHGGLQLFRRVGTELEAISGTLPPASQRWDRMLQASGRKGAKPSSASAQTSPQESVDAFADSLERELDAHSP MHQAGQTLASSRKRSRSESSVNRSSAQQAEEVFVPEQRDAPLLPLSSWGVRRRRTRIGGLPDPGTPTHGDLAASSAAFLEQDA DPFAGAAEDFPADFQEEIAWLKELLAH

Note: The 34-amino-acid repeats, due to variation in the 32nd-35th amino acids of repeats, are marked by gray color. The 12th and 13th amino acids of each repeat are variable and underlined. The absence of 13th amino acid in repeats is indicated by *.