

The potato NLR immune receptor R3a does not contain non-canonical integrated domains

Artemis Giannakopoulou, Angela Chaparro-Garcia, and Sophien Kamoun

The Sainsbury Laboratory, Norwich Research Park, Norwich, NR4 7UH, United Kingdom

A recent study by Kroj et al. (New Phytologist, 2016) surveyed nucleotide binding-leucine rich repeat (NLR) proteins from plant genomes for the presence of extraneous integrated domains that may serve as decoys or sensors for pathogen effectors. They reported that a FAM75 domain of unknown function occurs near the C-terminus of the potato late blight NLR protein R3a. Here, we investigated in detail the domain architecture of the R3a protein, its potato paralog R3b, and their tomato ortholog I2. We conclude that the R3a, R3b, and I2 proteins do not carry additional domains besides the classic NLR modules, and that the FAM75 domain match is likely a false positive among computationally predicted NLR-integrated domains.

Plants use immune receptors to recognize invading pathogens and pests. The largest family of intracellular immune receptors is the nucleotide binding-leucine rich repeat (NB-LRR or NLR) protein family - an important element of defense against pathogens in both plants and animals (Maekawa et al. 2011, Jacob et al. 2013). NLR proteins recognize pathogen-secreted molecules, termed effectors, which translocate into the host cytoplasm, triggering an immune response that leads to resistance (Jones and Dangl 2006, Ooijen et al. 2007, Dodds and Rathjen 2010, Win et al. 2012). NLR recognition of pathogen effectors can be direct through physical interaction between receptor and effector or indirect via receptor monitoring of effector-mediated modifications of plant target proteins (Mackey et al. 2002, Axtell and Staskawicz 2003, Rooney et al. 2005). Another model of NLR-detection of pathogen effectors has recently been described in which an extraneous domain, derived from a host target, has integrated into the NLR protein to enable effector recognition (Cesari et al. 2014, Nishimura et al. 2015, Wu et al. 2015). These “integrated domains” have been described as decoys, sensors, and baits, and appear diverse and widespread among plant NLR proteins (Cesari et al. 2014, Nishimura et al. 2015, Wu et al. 2015, Khan et al. 2016, Kroj et al. 2016, Sarris et al. 2016). To determine the extent to which NLR proteins carry integrated domains, Kroj et al (2016), surveyed NLRs from the genome sequences of 31 plant species for non-canonical domains. Among their predictions, they reported that a FAM75 domain of unknown function occurs near the C-terminus of the potato late blight resistance protein R3a (Kroj et al. 2016).

R3a is an NLR protein that confers resistance to the Irish potato famine pathogen *Phytophthora infestans* (Huang et al. 2005). It is one of the first *R* genes to be used in potato breeding for resistance against the late blight pathogen *Phytophthora infestans* (Armstrong et al. 2005, Bos et al. 2006, Morgan and Kamoun 2007, Birch et al. 2008). R3a carries a predicted coiled-coil (CC) domain in its N-terminal part (Huang et al. 2005) and shares 88% nucleotide identity and 83% amino acid similarity with I2, a tomato NLR that mediates resistance to the soil-borne fungus *Fusarium oxysporum f. sp. lycopersici* (Ori et al. 1997, Simons et al. 1998). The R3 locus includes another functional NLR gene, R3b, which shares 82% nucleotide

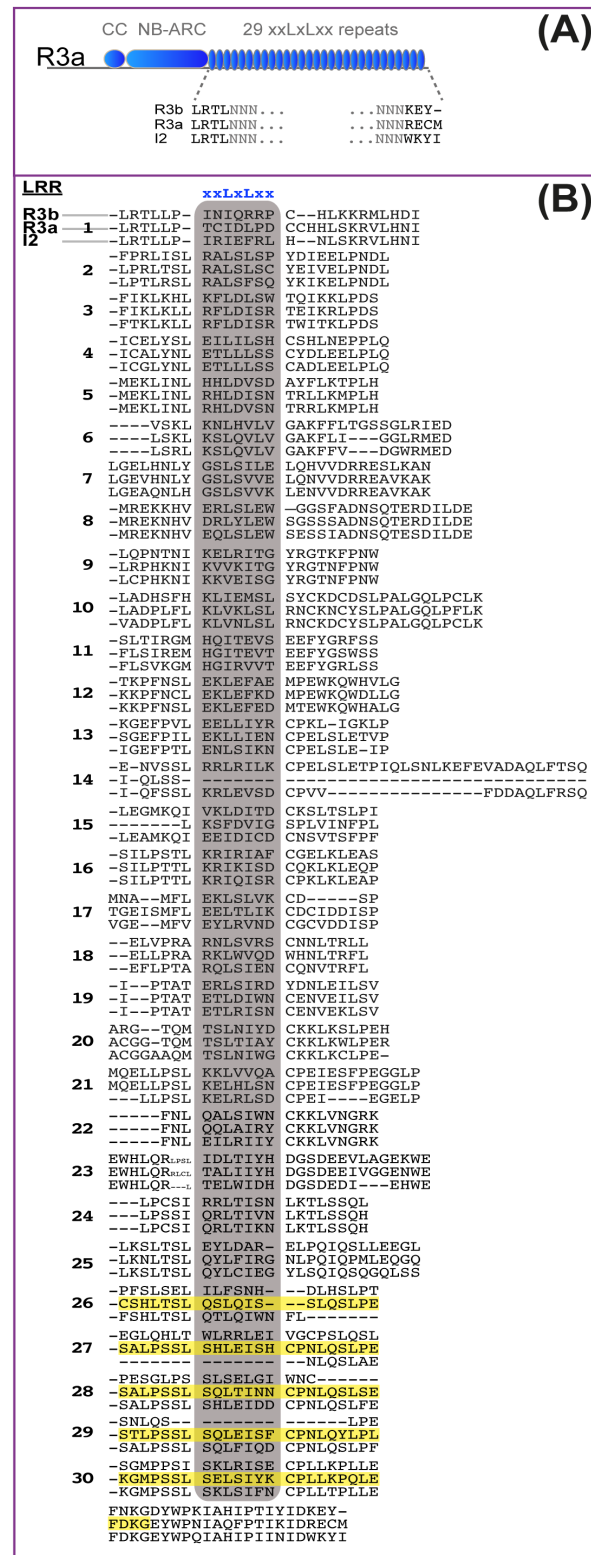
identity with R3a (Huang et al. 2004, Li et al. 2011). R3b also confers resistance to *P. infestans* but with a distinct recognition specificity compared to R3a (Huang et al. 2004, Li et al. 2011). R3a and R3b share their highest similarity in the CC domain (79% amino acid identity), whereas R3a and I2 are mostly similar in the nucleotide binding (NB) domain (86% amino acid identity) (Huang et al. 2005, Li et al. 2011).

In their survey of NLR-integrated domains, Kroj et al. (2016), reported that R3a contains a non-canonical domain, termed FAM75 (see Figure 1 of Kroj et al. 2016). This report surprised us given that R3a and other members of the R3/I2 superfamily are fairly compact NLRs with little room for extraneous domains in addition to the classic modules. Here, we revisited the domain architecture of R3a to check the robustness of this finding.

RESULTS AND DISCUSSION

To determine whether R3a carries an integrated domain in addition to classic NLR modules, we reexamined an alignment between R3a, its paralog R3b, and their tomato ortholog I2 (Figure 1). The three proteins diverge mainly in the leucine rich repeat (LRR) domain, yet all of them share 29 xxLxLxx repeat units (Figure 1A). Notably, the R3a region between amino acids (aa) 1148-1263, predicted to encode the FAM75 domain, clearly overlaps with five LRRs (Figure 1B).

Figure 1. R3a does not carry a FAM75 domain. **A**, Schematic diagram of the R3a protein. The LRR domains of R3b, R3a and I2 are shown. **B**, Alignment of the LRR domains of R3b (top), R3a (middle) and I2 (bottom). The LRR units in the alignment are numbered. The consensus xxLxLxx of each repeat unit is highlighted with a gray box. R3a aa residues 1148-1263 predicted by Kroj et al. (2016) to encode a FAM75 domain are highlighted in yellow. The following protein sequences were used: R3b (GenBank accession AEC47890.1) R3a (GenBank accession AAW48299.1) and I2 (GenBank accession ALF36835.1).



To further examine the domain architecture of R3a, we performed an InterPro protein sequence search (Mitchell et al. 2015). This failed to identify additional domains besides the LRR in the C-terminal half of R3a, consistent with the alignment and manual annotation of the LRR (Figure 2).

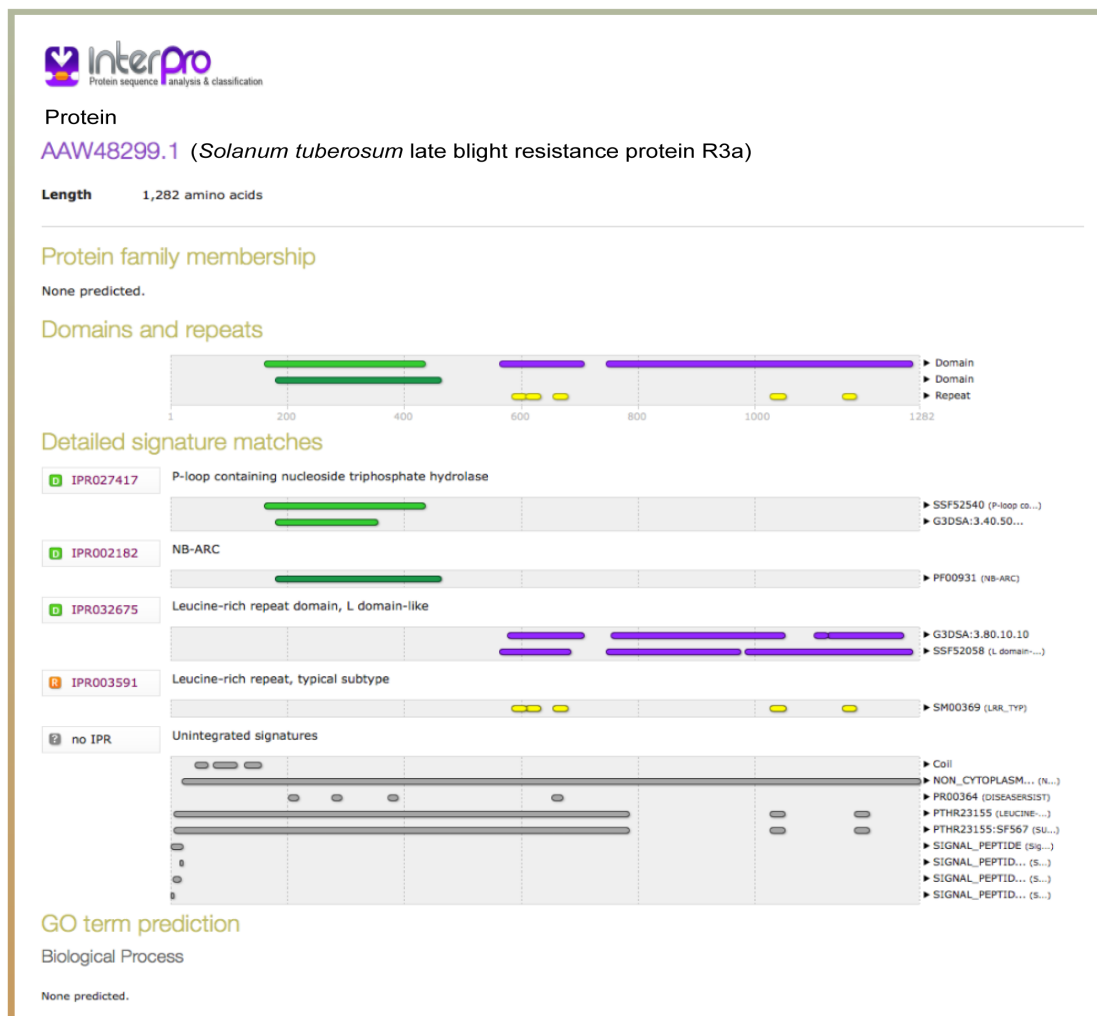
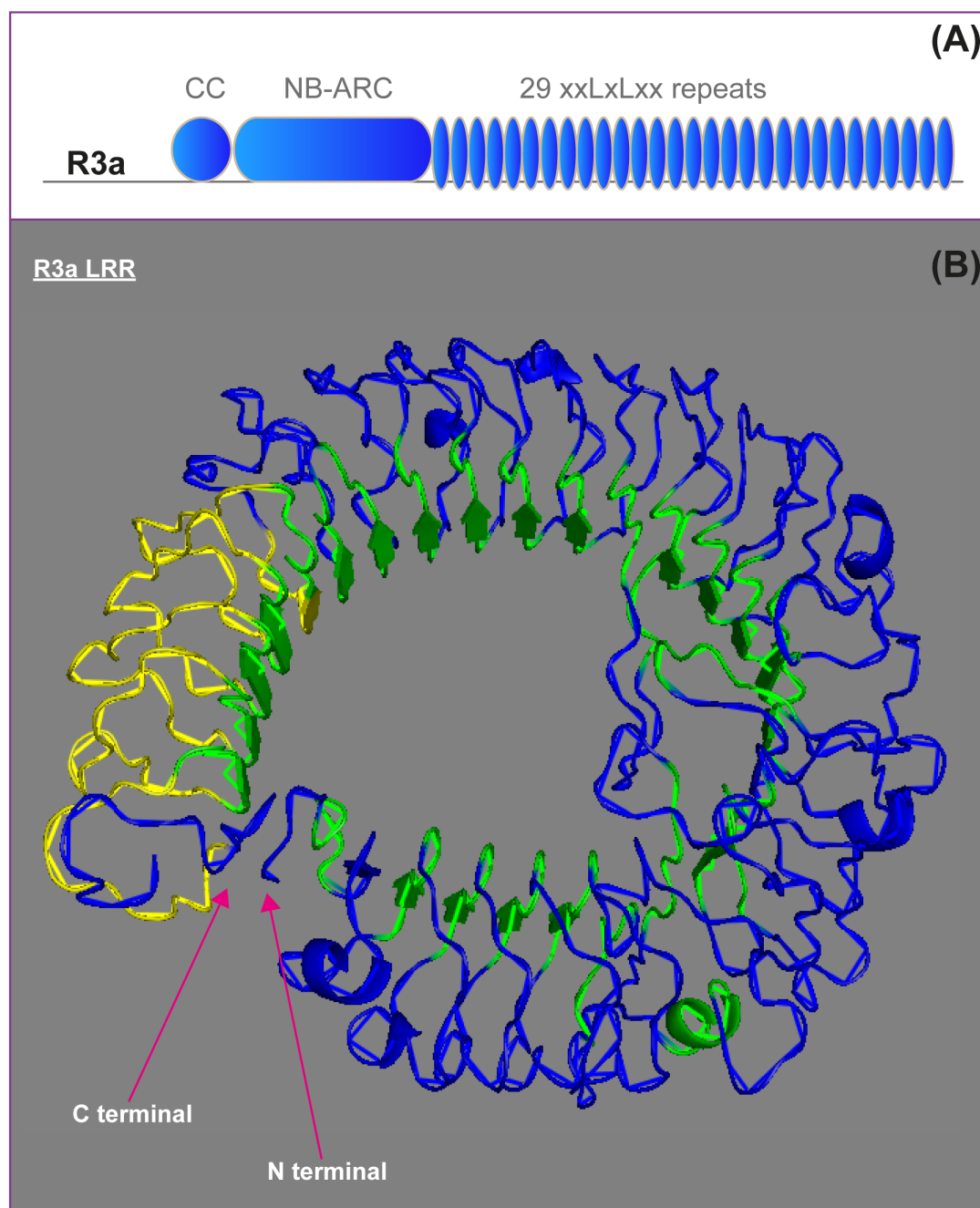


Figure 2. R3a protein domain architecture. We applied the InterPro protein sequence tool to the R3a sequence (Mitchell et al. 2015). Predicted domains and important sites of R3a are shown. The following R3a protein sequence was used for the analysis (GenBank accession AAW48299.1).

Next, we performed homology structure modeling of the R3a LRR region based on protein fold recognition algorithms implemented by IntFold (Figure 3) (Buenavista et al. 2012). We modeled the R3a LRR domain using as a template the top scoring Intfold protein, the Toll-like receptor 13 (PDB code 4z0c) (Song et al. 2015). The homology model confirmed that the presumed FAM75 region overlaps with the last five LRRs of R3a (Figure 3).

Figure 3 (next page). Structure homology model of the leucine rich repeat domain of R3a. **A.** Schematic diagram of the R3a protein. **B.** Structural representation of the R3a LRR domain. We generated the model by homology modeling using the IntFold server (Buenavista et al. 2012). LRRs are highlighted in green. The C-terminal region at amino acids 1148-1263 predicted by Kroj et al. (2016) to match the FAM75 domain is highlighted in yellow. The R3a protein sequence used for the analysis is GenBank accession AAW48299.1.



We conclude that R3a does not carry additional domains besides the classic CC, NB, and LRR modules of NLR proteins. Our analysis highlights the possibility of false positives among computationally predicted NLR-integrated domains. The position of predicted integrated domains needs to be checked for overlap with canonical NLR domains. Other false positives may include the DUF3542 domain, which overlaps with the CC domain of several solanaceous NLR proteins and may be an extension of the CC domain. DUF3542 was reported as the second most abundant integrated domain (Sarris et al. 2016), and was assigned to R1, another potato late blight resistance protein (Kroj et al. 2016). In R1 (GenBank accession AF447489_1), the CC domain (amino acids 406-515) overlaps with DUF3542 (amino acids 111-515) indicating that these two domains might be related.

ACKNOWLEDGEMENTS

Research in the Kamoun Lab is funded by the Biotechnology and Biological Sciences Research Council (BBSRC), the European Research Council (ERC), and the Gatsby Charitable Foundation.

REFERENCES

- Armstrong, M. R., et al. (2005). An ancestral oomycete locus contains late blight avirulence gene Avr3a, encoding a protein that is recognized in the host cytoplasm. *Proceedings of the National Academy of Sciences of the United States of America* **102**:7766-7771.
- Axtell, M. J. and B. J. Staskawicz (2003). Initiation of RPS2-specified disease resistance in *Arabidopsis* is coupled to the AvrRpt2-directed elimination of RIN4. *Cell* **112**:369-377.
- Birch, P. R. J., et al. (2008). Oomycete RXLR effectors: delivery, functional redundancy and durable disease resistance. *Current Opinion in Plant Biology* **11**:373-379.
- Bos, J. I. B., et al. (2006). The C-terminal half of *Phytophthora infestans* RXLR effector AVR3a is sufficient to trigger R3a-mediated hypersensitivity and suppress INF1-induced cell death in *Nicotiana benthamiana*. *Plant Journal* **48**:165-176.
- Buenavista, M. T., et al. (2012). Improvement of 3D protein models using multiple templates guided by single-template model quality assessment. *Bioinformatics* **28**:1851-1857.
- Cesari, S., et al. (2014). A novel conserved mechanism for plant NLR protein pairs: the "integrated decoy" hypothesis. *Frontiers in Plant Science* **5**:606-606.
- Dodds, P. N. and J. P. Rathjen (2010). Plant immunity: towards an integrated view of plant-pathogen interactions. *Nature Reviews Genetics* **11**:539-548.
- Huang, S., et al. (2004). The R3 resistance to *Phytophthora infestans* in potato is conferred by two closely linked R Genes with distinct specificities. *Molecular Plant-Microbe Interactions* **17**:428-435.
- Huang, S. W., et al. (2005). Comparative genomics enabled the isolation of the R3a late blight resistance gene in potato. *Plant Journal* **42**:251-261.
- Jacob, F., et al. (2013). Evolution and Conservation of Plant NLR Functions. *Frontiers in immunology* **4**:297-297.
- Jones, J. D. G. and J. L. Dangl (2006). The plant immune system. *Nature* **444**:323-329.
- Khan, M., et al. (2016). Of guards, decoys, baits and traps: pathogen perception in plants by type III effector sensors. *Current Opinion in Microbiology* **29**:49-55.
- Kroj, T., et al. (2016). Integration of decoy domains derived from protein targets of pathogen effectors into plant immune receptors is widespread. *New Phytologist* **210**:618-626.
- Li, G., et al. (2011). Cloning and characterization of R3b; members of the R3 superfamily of late blight resistance genes show sequence and functional divergence. *Molecular Plant-Microbe Interactions* **24**:1132-1142.
- Mackey, D., et al. (2002). RIN4 interacts with *Pseudomonas syringae* type III effector molecules and is required for RPM1-mediated resistance in *Arabidopsis*. *Cell* **108**:743-754.
- Maekawa, T., et al. (2011). Coiled-coil domain-dependent homodimerization of intracellular barley immune receptors defines a minimal functional module for triggering cell death. *Cell Host & Microbe* **9**:187-199.

Mitchell, A., et al. (2015). The InterPro protein families database: the classification resource after 15 years. *Nucleic Acids Research* **43**:D213-D221.

Morgan, W. and S. Kamoun (2007). RXLR effectors of plant pathogenic oomycetes. *Current Opinion in Microbiology* **10**:332-338.

Nishimura, Marc T., et al. (2015). Treasure your exceptions: unusual domains in immune receptors reveal host virulence targets." *Cell* **161**:957-960.

Ooijen, G. v., et al. (2007). Structure and function of resistance proteins in solanaceous plants. *Annual Review of Phytopathology* **45**:43-72.

Ori, N., et al. (1997). The I2C family from the wilt disease resistance locus I2 belongs to the nucleotide binding, leucine-rich repeat superfamily of plant resistance genes. *Plant Cell* **9**:521-532.

Rooney, H. C. E., et al. (2005). *Cladosporium* Avr2 inhibits tomato Rcr3 protease required for Cf-2-dependent disease resistance. *Science* **308**:1783-1786.

Sarris, P. F., et al. (2016). Comparative analysis of plant immune receptor architectures uncovers host proteins likely targeted by pathogens. *BMC Biology* **14**:1-19.

Simons, G., et al. (1998). Dissection of the *Fusarium* I2 gene cluster in tomato reveals six homologs and one active gene copy. *Plant Cell* **10**:1055-1068.

Song, W., et al. (2015). Structural basis for specific recognition of single-stranded RNA by Toll-like receptor 13. *Nature Structural Molecular Biology* **22**:782-787.

Win, J., et al. (2012). Effector biology of plant-associated organisms: concepts and perspectives. *Cold Spring Harbor Symposia on Quantitative Biology* **77**:235-247.

Wu, C.-H., et al. (2015). The "sensor domains" of plant NLR proteins: more than decoys? *Frontiers in Plant Science* **6**:134-134.