

PREreview of bioRxiv article “Subfamily-specific functionalization of diversified immune receptors in wild barley”

Sophien Kamoun¹

¹The Sainsbury Laboratory

April 28, 2020

Abstract

This is a review of Maekawa et al. bioRxiv 293050; doi: <https://doi.org/10.1101/352278> posted on June 20, 2018. In this paper, the authors mined the transcriptomes of 50 different accession of wild barley, generating a rich library of natural variants of the MLA immune receptor—a classical nucleotide-binding domain and leucine-rich repeat-containing (NLR) protein. They grouped the MLA variants in two subfamilies with all receptors known to be effective against the powdery mildew fungus grouping in one subfamily.

Summary

In plants, intracellular immune receptors of the NLR (nucleotide-binding, leucine-rich repeat) family are modular proteins that monitor translocated pathogen effector proteins and activate immune responses typified by the hypersensitive cell death. One of these NLR proteins is MLA, a member of a group of proteins that carry a coiled-coil (CC) domain prior to the nucleotide-binding (NB) and leucine-rich repeat (LRR) domains. In this paper, the authors mined the transcriptome of 50 different accession of wild barley, generating a rich library of natural variants of the MLA immune receptor. They found a pattern of diversification in the CC domain, which they argue might be related to functional diversification of these receptors. Furthermore, they detected positive selection signals in the LRR region of MLA, which is thought to confer recognition specificity to the pathogen.

The findings represent an excellent example of molecular evolution in plant NLRs and the new receptor variants uncovered by this study have the potential to help in the quest for durable resistance against different pathogen strains.

There are two parts in this study: (1) the generation of a sequence library of natural variants of the MLA gene and the analyses to identify signatures of selection and diversification; (2) a combination of secondary structure prediction and functional analyses of the CC domains to complement the first part.

Our general view is that the first part is interesting and has yielded exciting molecular evolution findings. Figure S2, for instance, is truly beautiful with MLA being fairly conserved in its general structure yet so diverse in terms of amino acid sequences. However, the second part would benefit from revising the methodology related to the structural analyses of the CC domains. Also, the functional analyses are limited to autoactivity assays of the CC-domains. It's not surprising that all of the assayed domains trigger cell

death given that the CC domains of wheat MLA-like genes Sr33 and Sr50, which group outside the two clades described here, are also autoactive. Given that all of the assayed domains are autoactive, and appear to be conserved across the broader MLA family including wheat homologs, it is difficult to accept several of the conclusions proposed in this manuscript. For example, the title (“subfamily-specific functionalization”) might be misleading given that no pathogen or effector related functional data on subfamily 2 has been reported.

Findings and comments

(1) Overall comments

The lack of functional analyses with pathogens and/or effectors and the fact that the study does not identify a differential phenotype between the reported subfamilies, do not support the statements about functional diversification.

CC domain differences are proposed to drive the diversification of the MLA receptor. However, the phylogeny of the conserved NB domain does not fully recapitulate this diversification.

The manuscript would benefit from a more thorough functional assessment of a range of CC domains. Only one CC domain from the new subclade, which is also the closest homolog to Sr50, has been assayed. Testing more members across the newly identified subfamily will help to draw a general conclusion about the new subclade.

The manuscript proposes a possible diversification in signalling capabilities, however, the chimera experiments of (Jordan et al., 2011) are not consistent with this hypothesis and suggest conserved signalling capacities.

(2) Comments on structural predictions

The 21st residue of the MLA family CC domains is generally occupied by an aspartate or a glutamate whereas it is a glycine in Sr33, and it is suggested in the manuscript that this may account for the reported differences in the structures—Sr33 was described as a four-helix bundle by (Casey et al., 2016), and MLA10 as a helix-loop-helix in an obligate dimer by (Maekawa et al., 2011). The (Casey et al., 2016) paper has shown that the CC domains of Sr33, MLA10 and Rx all maintain the same oligomeric state and four-helix bundle fold in solution, and this is supported by biophysical analyses of recombinantly produced protein. As such, the current debate on the CC domain structure is not centred around differences in their tertiary structures. What is unknown is whether the MLA10 CC domain dimeric helix-loop-helix structure represents an alternative quaternary conformation, for example a post-activation conformation. To this end, the manuscript does not address the “alternative activation state hypothesis” (discussed by (El Kasmi and Nishimura, 2016)), rather the text implies that the tertiary structures of Sr33 and MLA10 are different. To support the statement that Sr33 and MLA10 CC domains maintain different tertiary structures, the authors applied secondary structure prediction with PSIPRED and protein stability modelling with the STRUM web-server. However, published biochemical and biophysical data demonstrating the structural similarity of the Sr33 and MLA10 proteins in solution are not fully considered.

Secondary structure predictions of the MLA10 and Sr33 CC domains were stated to be performed with the first 40 amino acids “for simplicity”. This is problematic as secondary structure prediction using PSIPRED

can vary depending on the length of the sequence submitted. Indeed, when the first 160 amino acids of MLA10 and Sr33 are submitted to PSIPRED, the observed differences in the “looped vs helical” regions of the first 40 residues of Sr33 and MLA10 reported in this manuscript are no longer apparent. Considering that the expression of the 1-160 region of the CC domains (or equivalent) triggered cell death in planta (Figure 4a), we believe this region to be more appropriate for secondary structure predictions.

It was hypothesized that the presence of the glycine at the 21st residue in Sr33 is the determinant of the “structural differences between MLA10 and Sr33”, and subsequently used the STRUM server (structure-based prediction of protein stability changes upon single-point mutation), to predict whether reciprocal mutations of polymorphisms between MLA10 and Sr33 (Sr33 V20T and Sr33 G21E; MLA10 T20V and MLA10 E21G) would be sufficient to destabilise the MLA10 and Sr33 structures, respectively. The results of STRUM suggest the MLA10 T20V and MLA10 E21G would likely destabilise the MLA10 structure, however the reciprocal mutations in Sr33 would have no effect. There are several questions that this analysis raises listed below.

- What was the structure used to model the destabilisation? The original MLA10 structure (3QFL) could be compared to the four-helix bundle fold of the Rx and Sr33 CC domains. The small-angle X-ray scattering (SAXS) data for MLA10 CC domain published by (Casey et al., 2016) indicates the MLA10 CC domain also likely forms a four-helix bundle in solution, therefore a better approach to this experiment would be to generate a homology model of MLA10 based on the Sr33 CC domain four-helix bundle and then assess the effects of the mutants on protein stability using STRUM. Additionally, it would be ideal if all the approaches taken to the predictions in the manuscript could be detailed in the materials and methods.
- To simulate protein stability, it would be more appropriate to use the entire functional region of the protein instead of a non-functional shorter fragment, as it is likely that the missing residues could contribute to stability of the protein. Unfortunately, there is no structure of the active CC domain of MLA10 (a minimum of 142 residues) and as STRUM requires a structure to predict destabilisation caused by point mutations, it is not possible to determine the destabilising effect of the Sr33 V20T and Sr33 G21E mutants in a functional CC domain, even via homology modelling. Consequently, any stability prediction using either the current MLA10 or Sr33 CC domain structures are essentially not directly relevant to the minimal functional domain.
- Biochemical and biophysical analyses would be the more robust approaches than predictions of protein stability. These experiments are possible given that MLA10 CC domain can be purified in quantities sufficient for structural studies as in (Maekawa et al., 2011). For example, stability of MLA10 and Sr33 CC domain mutants could be analysed by circular dichroism (CD), 2D NMR, or by ThermoFluor. These experiments would be much more conclusive than prediction servers, and these data would significantly benefit the manuscript.
- Finally, the MLA10 T20V and MLA10 E21G mutants appear to not have any effects on the cell death phenotype nor on the accumulation of these proteins in planta (Fig. 5). These observations are inconsistent with the hypothesis that the mutations destabilize the proteins. Further discussion of these observations would be beneficial.

(3) Other comments

Figure S1 is a great positive control for the bioinformatics pipeline.

It is not clear why the second clade is described as a SUB-family of MLA given that ALL known MLA are in the other clade. The second clade is better described as MLA-like or MLA sister clade.

Some plants carry two (or even three) members of MLA. In these cases, do they belong to the same or a different subclade? It was not clear in the text and is worth commenting as this situation complicates allelic analyses.

Line 127: as there is not structural data available and to avoid confusion, the text should state “predicted to be located...”

Lines 142-143: The statement that RGH1/MLA family has been driven by subfamily-specific functionalization to distinct pathogens is highly speculative. Is there evidence for a second pathogen? Is possible that this subfamily detects uncharacterized powdery mildew strains. This contradicts lines 410-412 “Whether subfamily 2 NLRs confer disease resistance to avirulence genes present in yet uncharacterized Bgh populations or other pathogens remains to be tested”.

Line 257: “Bootstrap not very high”. Perhaps include the bootstrap number in brackets.

Line 384. It is unclear how they can conclude from RNAseq data only that “in wild barley Rgh1/Mla has undergone frequent gene duplication (Table S1)”. Could these sequences be allelic?

Lines 450-456. They cite (Shen et al., 2007) as evidence that MLA-CC functions by binding WRKY transcription factors and derepressing them. Our understanding is that this model was drawn from an experiment with the inappropriate avirulence effector.

Figures 4 and 5: the loading control would be easier to distinguish when showing the band corresponding to RuBisCO (55KDa).

Reviewers

Adam R. Bentham and Juan Carlos De la Concepcion, Department of Biological Chemistry, John Innes Centre, Norwich Research Park, Norwich, UK.

Sophien Kamoun. The Sainsbury Laboratory, Norwich Research Park, Norwich, UK.

References

- LW Casey, P Lavrencic, AR Bentham, S Cesari, DJ Ericsson, T Croll, D Turk, PA Anderson, AE Mark, PN Dodds, M Mobli, B Kobe, and SJ Williams. The CC domain structure from the wheat stem rust resistance protein Sr33 challenges paradigms for dimerization in plant NLR proteins. *Proc Natl Acad Sci U S A*, Oct 2016.
- F El Kasmi and MT Nishimura. Structural insights into plant NLR immune receptor function. *Proc Natl Acad Sci U S A*, Nov 2016.

- T Jordan, S Seeholzer, S Schwizer, A Töller, IE Somssich, and B Keller. The wheat Mla homologue TmMla1 exhibits an evolutionarily conserved function against powdery mildew in both wheat and barley. *Plant J*, 65:610–21, Feb 2011.
- T Maekawa, W Cheng, LN Spiridon, A Töller, E Lukasik, Y Saijo, P Liu, QH Shen, MA Micluta, IE Somssich, FLW Takken, AJ Petrescu, J Chai, and P Schulze-Lefert. 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, Mar 2011.
- QH Shen, Y Saijo, S Mauch, C Biskup, S Bieri, B Keller, H Seki, B Ulker, IE Somssich, and P Schulze-Lefert. Nuclear activity of MLA immune receptors links isolate-specific and basal disease-resistance responses. *Science*, 315:1098–103, Feb 2007.