

University of Dundee

DOCTOR OF PHILOSOPHY

**Computational Identification and Functional Characterisation of Candidate DNA Binding Effectors in Phytophthora**

Motion, Graham B.

*Award date:*  
2015

*Awarding institution:*  
University of Dundee

[Link to publication](#)

**General rights**

Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the public portal

**Take down policy**

If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

**Supplementary File 4:** Instructions for use of our plant specific prediction model in WEKA.

1. Download and install the latest developer version of WEKA for your operating system (<http://www.cs.waikato.ac.nz/ml/weka/downloading.html>).
2. Using the Perl script provided (Additional data file 6) convert your fasta file of plant amino acid sequences to be predicted to .arff format for input into WEKA.
3. Open the WEKA explorer and open the .arff file created in step 2.
4. Select the classify tab and check “supplied test set” under the test options. Set this as your input .arff file by clicking “set...” -> “open file...” -> select your file -> close.
5. Click on “more options...” and click on “choose” next to output predictions to select the format you would like your output predictions to appear in.
6. Right click on the white result list box and select load model.
7. Open the plant lineage specific model (Additional data file 5).
8. Right click on the model in the results list and select “Re-evaluate model on current test set”. This will run the prediction model on your input dataset.
9. Right click on the model in the results list for the option to save the results buffer containing which proteins are predicted as being DNA binding.