

Contents

	page
Abstract	v
Acknowledgements	vii
Contents	ix
List of Figures	xi
List of Tables	xv
List of Symbols	xvii
List of Abbreviations	xix
1 Introduction	1
1.1 Agriculture and sustainable developmental goals	1
1.2 Role of <i>Oryza sativa</i> in achieving sustainable developmental goals	1
1.3 Functioning of immune system of <i>Oryza sativa</i> against biotic and abiotic stresses	3
1.4 Current trends in enhancing stress-tolerance in <i>Oryza sativa</i>	5
1.5 Thesis statement	7
1.6 Scope of thesis	8
2 Literature review	11
2.1 Integrative breeding strategies in plant stress tolerance	11
2.1.1 Conventional plant breeding approaches	11
2.1.2 Modern plant breeding approaches	13
2.2 Advanced computational approaches to uncovering stress tolerance mechanisms in <i>Oryza sativa</i>	16
2.2.1 <i>In silico</i> characterization of novel disease-resistance proteins in <i>Oryza sativa Japonica</i> against bacterial leaf blight and rice blast diseases	16
2.2.2 Deep learning-based identification of disease-resistance proteins in <i>Oryza sativa</i> and its related species	17
2.2.3 AP2/ERF transcription factor orthologs of the desert tree <i>Prosopis cineraria</i> show higher copy number and DNA-binding affinity than drought-sensitive species	17
2.3 Conclusion	18
3 Characterization of novel disease-resistance proteins in <i>Oryza sativa Japonica</i> against bacterial leaf blight and rice blast diseases	19
3.1 Introduction	19
3.2 Materials and Methods	20
3.2.1 Data collection	20
3.2.2 Gene network construction and functional analysis	20
3.2.3 Physicochemical analysis of unknown disease-resistance proteins	21
3.2.4 Structure prediction and stereochemical property analysis of unknown disease-resistance proteins	21
3.2.5 Subcellular localization analysis	22
3.2.6 Multiple sequence alignment and phylogenetic analysis	22
3.2.7 Functional enrichment analysis of disease-resistance proteins	23
3.2.8 Identification of functional interactors in disease-resistance proteins	23
3.2.9 <i>In silico</i> gene expression analysis	24
3.3 Results	25
3.3.1 Identification of unknown disease-resistance genes in <i>Oryza sativa Japonica</i> through gene network analysis	25
3.3.2 Physicochemical characteristics of unknown <i>Oryza sativa Japonica</i> disease-resistance proteins	26
3.3.3 Three-dimensional structure of unknown disease-resistance proteins in <i>Oryza sativa Japonica</i>	27
3.3.4 Stereochemical features of unknown disease-resistance proteins in <i>Oryza sativa Japonica</i>	28
3.3.5 Subcellular localization of unknown disease-resistance proteins in <i>Oryza sativa Japonica</i>	29
3.3.6 Phylogeny of unknown disease-resistance proteins in <i>Oryza sativa Japonica</i>	31
3.3.7 Predicted functions of unknown disease-resistance proteins in <i>Oryza sativa Japonica</i>	34
3.3.8 Predicted functional interactors of unknown disease-resistance proteins in <i>Oryza sativa Japonica</i>	35
3.3.9 Gene expression under biotic stress of unknown disease-resistance proteins in <i>Oryza sativa Japonica</i>	36
3.4 Discussion	37

4	Deep learning based prediction of disease-resistance proteins in <i>Oryza sativa</i>	41
4.1	Introduction	41
4.2	Materials and Methods	42
4.2.1	Data collection	42
4.2.2	Data preprocessing	43
4.2.3	Feature extraction	43
4.2.4	Feature selection	46
4.2.5	Modeling of disease-resistance protein on the internal dataset	47
4.2.6	Evaluation of the DL-based model with ML models on an external dataset	50
4.2.7	Comparison of performance of the MLP model with well-established models	50
4.3	Results	51
4.3.1	Data optimization for model performance	51
4.3.2	Protein feature extraction and selection	52
4.3.3	Evaluation of the ML and DL model performance on the internal dataset	53
4.3.4	Evaluation of the ML and DL model performance on the external dataset	55
4.3.5	Comparison of the DL-based MLP model performance with existing models to predict disease-resistance proteins	56
4.3.6	Discussion	57
4.3.7	Conclusion	59
5	Genome-wide transcriptomic studies of the desert tree <i>Prosopis cineraria</i> to uncover drought tolerance genes	61
5.1	Introduction	61
5.2	Materials and methods	62
5.2.1	Reference-guided transcriptome analysis of <i>Prosopis cineraria</i> under drought stress	62
5.2.2	<i>De novo</i> transcriptome assembly of the Indian cultivar of <i>Prosopis cineraria</i>	63
5.2.3	Network analysis of <i>Prosopis cineraria</i> genes differentially expressed under drought	63
5.2.4	Phylogenetic classification of AP2/ERF proteins in <i>Prosopis cineraria</i> and <i>Oryza sativa Japonica</i> and <i>Indica</i>	64
5.2.5	Real-time quantitative PCR analysis of <i>Prosopis cineraria</i> AP2/ERF genes under drought	64
5.2.6	Analysis of the conserved motifs and gene structures in <i>Prosopis cineraria</i> AP2/ERF proteins	64
5.2.7	Promoter and <i>cis</i> -acting element analysis of AP2/ERF superfamily genes in <i>Prosopis cineraria</i>	65
5.2.8	Polymorphism analysis in AP2/ERF genes between Indian and Arabian cultivars of <i>Prosopis cineraria</i>	65
5.2.9	Copy-number variation analysis of the AP2/ERF superfamily between drought-tolerant and sensitive species	65
5.2.10	Modeling of apo and DNA-bound structures of the AP2/ERF and DREB proteins from drought-tolerant and sensitive species	65
5.2.11	Analysis of hydropathy and electrostatic properties of AP2/ERF proteins	66
5.2.12	All-atom molecular dynamics simulations	66
5.2.13	Hydrogen bond analysis	66
5.2.14	Binding energy calculation	66
5.2.15	3D structure and <i>in silico</i> DNA binding analysis of AP2/ERF proteins of <i>Prosopis cineraria</i> , <i>Oryza sativa Japonica</i> and <i>Indica</i>	67
5.3	Results	67
5.3.1	Stress-tolerance genes are highly induced in <i>Prosopis cineraria</i> under drought	67
5.3.2	Classification of <i>Prosopis cineraria</i> AP2/ERF genes	70
5.3.3	<i>Prosopis cineraria</i> AP2/ERF genes display different expression patterns under drought stress	70
5.3.4	<i>Prosopis cineraria</i> AP2/ERF promoters contain <i>cis</i> -elements for abiotic stress response	71
5.3.5	Gene structure and conserved motifs of <i>Prosopis cineraria</i> AP2/ERF proteins	73
5.3.6	AP2/ERF gene polymorphisms between Indian and Arabian cultivars of <i>Prosopis cineraria</i>	74
5.3.7	Drought-tolerant legumes contain more copies of AP2/ERF genes than drought-sensitive plants	81
5.3.8	The DNA-binding domains of <i>Prosopis cineraria</i> and <i>Arabidopsis thaliana</i> ERFs adopt the same fold but have chemical dissimilarity at a few positions	81
5.3.9	Differences in hydrogen bonding profile and charge distribution at the N- and C-terminal tails between <i>Prosopis cineraria</i> and <i>Arabidopsis thaliana</i> ERFs might be crucial in their differential DNA search and binding	83
5.3.10	DREB proteins have similar structures in drought-tolerant and drought-sensitive species	84
5.3.11	Classification of AP2/ERF genes in <i>Oryza sativa Japonica</i> and <i>Indica</i>	84
5.3.12	Structural differences between the AP2/ERF proteins of <i>Prosopis cineraria</i> , <i>Oryza sativa Japonica</i> and <i>Indica</i>	86
5.4	Discussion	90
6	Conclusions and Future work	95
6.1	Conclusions	95
6.2	Scope of Future Work	96
	Bibliography	99

Appendices	113
A Supplementary information for Chapter 3	115
B Supplementary information for Chapter 4	125
C Supplementary information for Chapter 5	131
Publications	135