

Mining and characterization of *Allium cepa* expressed sequence tags (ESTs) encoding receptor-like kinases (RLKs)

Satyabrata Nanda^{1,2}

¹Centre for Biotechnology, Siksha 'O' Anusandhan University, Bhubaneswar, India

²State Key Laboratory of Rice Biology, China National Rice Research Institute, Hangzhou, P.R. China

*Corresponding author: sbn.satyananda@gmail.com

Abstract

Plant receptor-like kinases (RLKs) are comprised of a large family of kinases having multitudinous functions. Onion (*Allium cepa*), an important bulbous plant, has almost no reports on RLKs and their role in onion plant growth and defense responses. In the present work, mining of the transcript database 'dbEST' has been performed to obtain the expressed tags of *A. cepa*. Subsequently, the consensus sequence assembling carried out by CAP3 program resulted in 4175 number of contigs. Consecutive rounds of BLAST searches resulted in identification of 4 highly homologous contigs with the previously reported plant RLKs. Further structural characterizations by MEME, Prosite and SOPMA supported the identified contigs to be putative RLKs. All four putative RLKs (AcCon13, AcCon183, AcCon383, and AcCon1738) possessed 9 out of 11 canonical kinase sub-domains, the signature protein-kinase domain, and the ATP binding sites. Phylogenetic analysis along with previously reported plant RLKs revealed that the identified contigs belong to different RLK sub-families.

Key words: dbEST, Receptor-like kinases, Contigs, *Allium cepa*

Abbreviation: ABA_Abscisic acid, EST_Expressed sequence tags, ORF_Open reading frame, RLK_Receptor-like kinase, STK_serine/threonine-protein kinase

Introduction

Plant Receptor-like kinase (RLK) is a multi-gene super-family involved in various physiological processes including plant growth and development, symbiosis, self-incompatibility, stress responses, and phytohormonal signaling (Gish and Clark, 2011; Antolin-Llovera et al., 2014). The plant RLKs typically consist of a signal sequence, a transmembrane domain, and a serine/threonine kinase domain (Greeff et al., 2012). A large number of RLKs are encoded by plants to perceive various internal and external stimuli. In *Arabidopsis*, more than 600 RLK members are been encoded which are involved in numerous cellular processes (Shiu et al., 2004). Besides their contributions in plant physiological processes, RLKs play a vital role in plant defense responses and pathogen recognition. The RLKs involved in plant defenses are known as pattern-recognition receptors (PRRs) which recognize the pathogen-associated molecular patterns (PAMPs) or microbe-associated molecular patterns (MAMPs) and instigate PAMP-triggered immunity (PTI) (Mengiste, 2012). Although, plant RLKs can be divided into as many as 44 families (Greeff et al., 2012), the best studied families are leucine-rich repeat (LRR), proline-rich extensin-like receptor protein kinase or proline-rich receptor-like protein kinase (PERK), lectin receptor-like kinase or lectin receptor kinase (LRK), LysM receptor-like kinases (LySM), and serine/threonine-protein kinase (STK). Recognition of the pathogen signatures by the RLKs result in a solitary or a complex cocktail of defense responses via generation of

reactive oxygen species (ROS), liberation of secondary messengers, synthesis of secondary metabolites, accumulation of phytohormones and expression of defense responsive genes (Andolfo and Ercolano, 2015).

Onion (*Allium cepa*) (family: Amaryllidaceae) is an important and most widely cultivated bulbous vegetable in the world. India is the second largest producer and the largest exporter of onions producing more than 19 million tonnes, and exporting more than 1 million tonnes per year, respectively (FAOSTAT, 2017). Although, the demand of onion ever rising, the vegetable faces huge crop losses due to pest and pathogen attacks majorly from insects, nematodes, and fungi. However, apart from some reports on markers associated with physiological traits and marker-assisted selection (MAS), no substantial research has been carried out to identify and isolate disease resistance genes in onions. Thus, identification and characterization of potential RLKs in onion will be highly beneficial in understanding the dynamics of onion resistance responses. Advancements in the omics approaches, and increasing number of next-generation sequencing (NGS) and microarray experiments have resulted in deposition of innumerable expressed sequence tags (ESTs) in public database like dbEST. The availability of these expressed tags facilitates the mining, prediction and characterization of candidate genes by following stringent bioinformatics procedures. Mining of available ESTs and genome survey sequences in the public

databases (dbESTs and dbGSS) have been used to identify numerous candidate genes including resistance gene homologs (RGH) (Ren et al., 2014), embryonic seed development genes (Emila and Percy, 2015), glycine-rich proteins (Kar et al., 2012), mitogen activated protein kinases (Joshi et al., 2011a), nucleotide binding site-leucine rich repeat (NBS-LRR) genes (Joshi et al., 2011b), microRNAs (Chand et al., 2017) and micro-satellites markers (Lopez et al., 2015; Chand et al., 2015). In the current work, an EST mining-based identification of putative RLKs in *A. cepa* has been carried out by using already reported RLK sequences as bait. Further, the structural characterization of these identified highly homologous contigs has been performed by using several bioinformatics tools.

Results and Discussion

Contig delineation by sequence similarity and sequence comparison

By using CAP3 sequence assembly program, a total of 4175 contigs were generated from 20225 *A. cepa* EST sequences. TBLASTN was used to perform reverse alignment of the generated contigs with some previously reported plant RLKs. The bait RLK sequences comprised of different RLK family sequences such as proline-rich receptor-like protein kinases (accession no. AQK94017.1, Q9LS95.2, KHF99092.1), lectin receptor kinase (accession no. ADK62373.1, AEE79957.1, AKV93688.1), LRR receptor-like protein kinase (accession no. AAF66615.1, KEH21564.1, BAE99117.1), receptor-like protein kinase (accession no. AAA32858.1, OAY72154.1, XP_002298383.2), serine/threonine-protein kinase (accession no. NP_001235164.1, AAV25281.1, XP_002304664.2), and LysM receptor-like kinase (Accession no. CAL46258.1, EEE80883.1, XP_002880525.1). Contigs which showed higher than 40% homology with the bait sequences were selected and used for another round of BLAST search. The second round of BLAST resulted in identification of four highly homologous contigs AcCon13, AcCon183, AcCon383, and AcCon1738 with significant hits and query coverage (score >400, query coverage >85%). AcCon183 and AcCon1738 showed the highest homology with serine/threonine-protein kinases from *Ananas comosus* and *Glycine max*, respectively. AcCon13 exhibited 80% similarity with PERK1 from *Zea mays*, whereas AcCon383 showed 85% similarity with Hercules receptor kinase1 (HERK1) from *Ananas comosus* (Table 1). Multiple sequence alignment (MSA) of AcCon13, AcCon1738, AcCon183, and AcCon383 revealed the presence of signature protein kinase sub-domains in all four contigs (Fig 1). Typically, a protein kinase consists of 11 canonical sub-domains, however, an additional signal peptide and a common docking (CD) domain can also be found at times (Hanks and Quinn 1991; Joshi et al., 2011a). The MSA results revealed that all four onion contigs possessed 9 out of 11 canonical sub-domains of a protein kinase with no signal peptide or CD domain.

Motif elicitation and structural characterization of the *A. cepa* contigs

De novo motif prediction by using multiple expectation maximization for motif elicitation (MEME) tool revealed the presence of 10 structural motifs in the four selected onion

contigs. Out of the 10 predicted *de novo* motifs, 9 motifs corresponded to I-IX domains of a protein kinase (Fig 2). Additional domain scans using Pfam and Prosite resulted in the identification of protein kinase signature domain and ATP-binding domain in all four contigs (Fig 3). The identification of protein kinase signature domains, and presence of the nine protein kinase sub-domains strongly suggested that the contigs belong to protein kinase family. A typical protein kinases functions by using a γ -phosphate of adenosine triphosphate (ATP) to phosphorylate an amino acid residue during phosphorylation process (Lehti-Shiu and Shiu, 2012). Thus, the detection of an ATP-binding region along with kinase domains in all four contigs confirmed their possible kinase roles. However, absence of two more kinase sub-domains and the average length of the four contigs (284 aa) suggests that the contigs might be the partial portions of potential RLKs. Translated sequences of all four onion contigs were further analyzed to predict the peptide properties. SOPMA analysis of the contigs revealed that all four *A. cepa* contigs constitute of majorly random coils (cc) (41.3% in AcCon13; 35.5% in AcCon183; 42.9% in AcCon1738; 37.6% in AcCon383) and α -helices (33.2% in AcCon13; 37.3% in AcCon183; 22.6% in AcCon1738; 24.4% in AcCon383) with extended stands and β -turns (Fig 4A). Similar kind of peptide properties were predicted for CLMPK6, a protein kinase involved in fungal defense in turmeric (Nanda et al., 2016). Hydropathy analysis of these four contigs was performed by using the Kyte-Doolittle plot. The results revealed that AcCon13, AcCon183, AcCon383 and AcCon1738 are hydrophobic in nature as they showed maximum peaks towards the negative side on the plot (Fig 4B).

Phylogenetic analysis and classification of *A. cepa* contigs

Phylogenetic analysis was performed to determine the relationship within the four *A. cepa* contigs and among the RLKs from other plant species. A neighbor-joining (NJ) tree was constructed based on the sequence information of the four onion contigs and 19 RLKs from other plants species. The resulted phylogenetic tree classified the onion contigs into three different RLK sub-families. AcCon13 formed a sub-clade with the PERKs, whereas Conting383 was grouped with the receptor-like protein kinases. AcCon183 and AcCon1738 both were grouped with the serine/threonine-protein kinases (Fig 5). AcCon383 showed 85% similarity with the RLK HERK1 from *A. comosus* and 77% similarity with the protein kinase from *Populus trichocarpa*. In *Arabidopsis*, HERK1 and HERK2 are reported to be involved in plant growth and stem elongation (Guo et al., 2009a; 2009b). Thus, AcCon383 being highly homologous to HERK1 might possibly be involved in regulating the plant growth and proliferation in onions. However, full-length characterization and in-depth molecular analysis of AcCon383 will help to confirm its role in cell growth. Similarly, AcCon13 shared 80% similarity with PERK1 from *Zea mays* and formed a sub-clade along with PERKs from *Arabidopsis* and *Gossypium*. Nakhmchik et al. (2004) has reported the *Arabidopsis* PERK family to be consisting of 15 predicted kinases. Further, the expressions of these PERKs in *Arabidopsis* were found to be both tissue-specific and generic. AtPERK4 was reported to be involved in abscisic acid (ABA) signaling response and root cell elongation (Bai et al., 2009). In *Brassica napus*, PERK1

Table 1. BLAST resulted sequences with significant homology to *A. cepa* contigs.

Contig name	Homologous sequence	Identity	Accession no.	E value
AcCon13	Proline-rich receptor-like protein kinase PERK1 [Zea mays]	80%	AQK94017.1	6e-162
AcCon1738	serine/threonine-protein kinase [Glycine max]	84%	NP_001235164.1	2e-145
AcCon183	Serine/threonine-protein kinase [Ananas comosus]	87%	OAY85824.1	2e-173
AcCon383	Receptor-like protein kinase HERK 1 [Ananas comosus]	85%	OAY72154.1	2e-144



Fig 1. Multiple sequence alignment of the predicted protein sequences of four *A. cepa* contigs analyzed by Clustal Omega. The characteristic subdomains of protein kinases are indicated by the roman numerals and the span of each subdomain is highlighted in grey. Asterisk (*) represents positions with fully conserved amino acid residue. Colon (:) represents positions conserved with similar amino acids. Period (.) represents positions conserved with weakly similar amino acids.

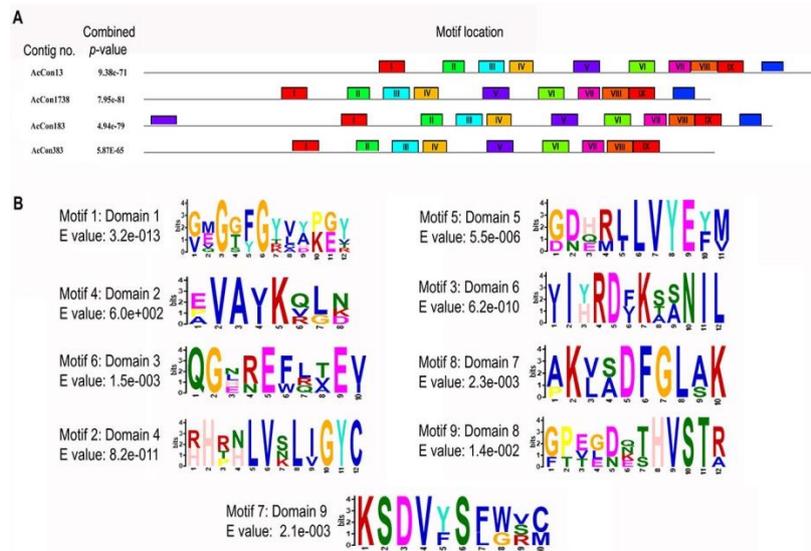


Fig 2. Prediction of de novo motifs from four *A. cepa* contig sequences by MEME tool. **A)** The solid line denotes the contig length whereas the colored boxes represent the predicted motifs along the length of each contig. **B)** The sequence logos along with the E-values of the predicted motifs from four contigs that corresponded to the conserved kinase subdomains.

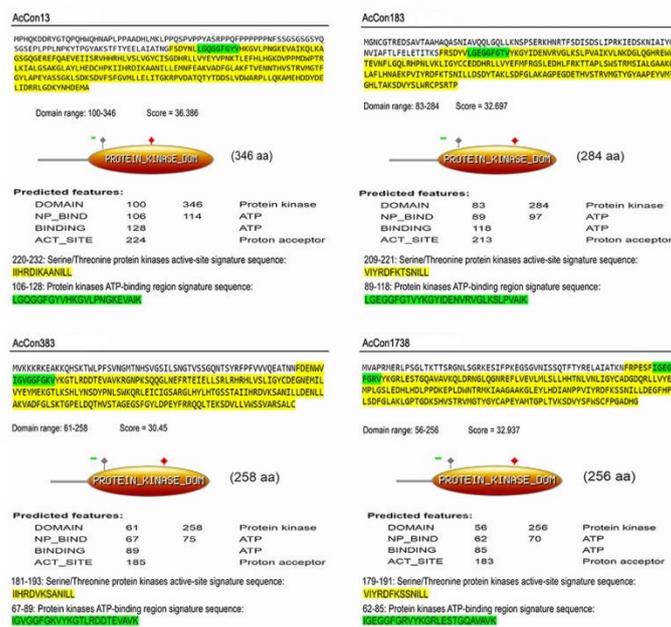


Fig 3: Prosite domain scan result of four *A. cepa* contigs showing the signature protein kinase domain and the ATP binding site.

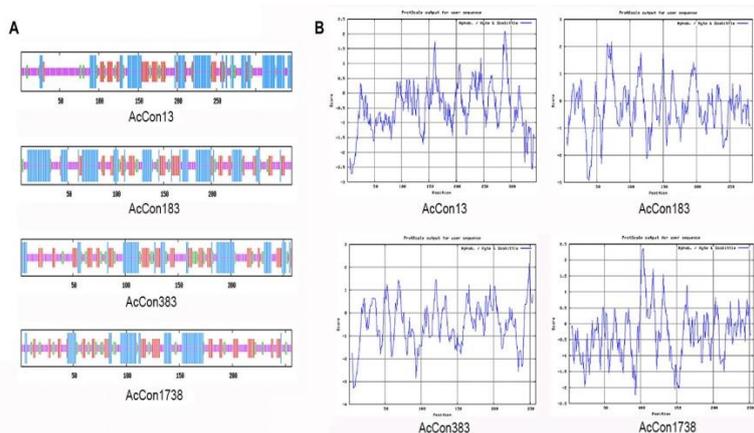


Fig 4. (A) Structural analysis of the deduced protein sequences of four *A. cepa* contigs using SOPMA. (B) Hydropathy analysis of the deduced protein sequences of four *A. cepa* contigs by Kyte-Doolittle plot. Y-axis shows the hydropathy scores whereas in X-axis amino acid length is covered. Stronger negative sided peaks suggest the hydrophobic nature of all four contigs.

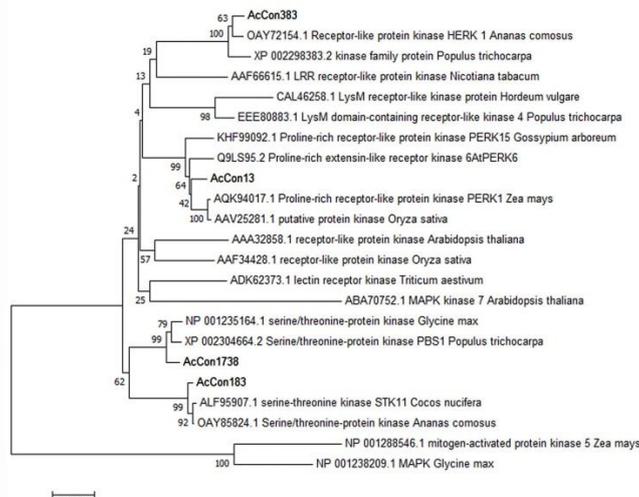


Fig 5. The phylogenetic analysis of four *A. cepa* contigs along with other reported plant RLKs constructed using MEGA 7.0. The accession numbers of the RLKs used to construct the tree is presented along with the RLK names. Numbers on the branches indicate the percentage of 1000 bootstrap replications supporting the particular nodes.

was reported to be involved in the early perception of mechanical injury and fungal attack (Silva and Goring, 2002). Thus, AcCon13, exhibiting high sequence homology and being grouped with PERKs, might have a similar role in onion as like the reported PERKs. AcCon1738 showed 84% homology with serine/threonine-protein kinase from *G. max* and 82% similarity with serine/threonine-protein kinases PBS1 from *P. trichocarpa*. Similarly, AcCon183 showed 87% homology to serine/threonine-protein kinase STK11 from *Cocos nucifera* and serine/threonine-protein kinases *A. comosus*. Conventionally, the protein kinases are divided into two families such as protein tyrosine kinases and serine/threonine kinases (Hanks and Quinn, 1991). The presence of comparatively poorly conserved GT/XXY/X-APE consensus than the highly conserved tyrosine-specific consensus PI/VK/RWT/MAPE in sub-domain VIII in both AcCon183 and AcCon1738 further confirms them to be STKs (Rudrabhatla et al., 2006). Plants STKs are one of the extensively studied protein kinase sub-family and have been reported to be involved in numerous cellular functions including plant defense (Afzal et al., 2008; Ho, 2015). Furthermore, STKs contribute to plant defense by regulating several physiological processes like cell wall strengthening, synthesis of secondary metabolites and inducing defense gene responses (Amil-Ruiz et al., 2011; Sekhwal et al., 2015). SnRK2 protein kinases, another family member of plant STKs reported to be involved in ABA signaling and abiotic stress responses (Kulik et al., 2011). Interestingly, STKs are also associated with negative regulation of plant defense. In *Arabidopsis*, PBL13, a STK is reported to negatively regulate defense against pathogenic bacteria *Pseudomonas syringae* attack (Lin et al., 2015). Having the serine/threonine kinase signature domains and being grouped with the reported plant STKs, AcCon183 and AcCon1738 claim to be potential STKs in onion which might have plant defense regulatory roles. Extensive structural and functional characterization of these two contigs will pave new insights into their possible involvement in onion defense responses. Although, further functional analysis is required to establish the physiological roles of the identified putative RLKs in onions, this work can serve as a basic framework for future research on delineating the diverse roles of *A. cepa* RLKs. In several previously reported studies, the *in silico* data mining from different databases have resulted in identification and characterization of a number of important genes in plants. By EST mining and from the available full-length cDNA sequences in public databases, Keeling et al. (2011) reported the phylogenetic classification and functional characterization of terpene synthase genes in spruce. Joshi et al (2011a) reported the identification of four MAPK partial sequences through the *in silico* analysis of the *Curcuma longa* EST database. This work later paved the way for cloning, isolation, and structural and functional characterizations of four MAPKs in *C. longa* under various abiotic stresses (Nanda et al., 2014). Data mining using HMMER and BLAST, and sequence retrievals from EST database had resulted in identification and characterization of Calmodulin-binding transcription activator (CAMTA) genes in strawberry (Leng et al., 2015). Apart from plants, mining of EST database and further functional analysis resulted in identification of effector proteins in the phytopathogen *Phytophthora infestans* (Torto et al., 2003). Similarly, *in silico* mining has resulted in characterization of 12 putative genes

encoding chitinase-like proteins in *Nilaparvata lugens*, a dreadful pest of rice (Xi et al., 2014). Thus, these current findings will help in the subsequent structural and functional characterization of the identified onion contigs, which in turn will reveal their actual roles in plant growth and defense responses.

Materials and methods

Sequence retrieval and processing from dbEST

The transcript database dbEST from NCBI is a public database that contains the expressed sequence tags or single-pass cDNA sequences of plants, animals, and micro-organisms. For the current *in silico* data mining work, transcript sequences of onion expressed in different physiological conditions and in different tissues were obtained from dbEST. A total of 20255 number of ESTs were retrieved in FASTA format and used for further processing. As the sequences often contain unwanted sequence contaminations such as part of vector sequences or adapter sequences, all 20255 ESTs were processed to obtain clean reads. The EST sequences were screened against the UniVec database of NCBI (<ftp://ftp.ncbi.nlm.nih.gov/pub/UniVec/>) to detect vector and adapter sequence contaminations and such detected sequences were subsequently removed. Clean reads with no sequence contaminations were used subsequently for generation of contigs.

Generation of contigs and homology search

CAP3 sequence assembly program (Huang and Madan, 1999) was employed to generate contigs and to create a non-redundant dataset from the cleaned EST sequences. The Basic Local Alignment Search Tool (BLAST) variant TBLASTN was used to perform reverse alignment of some reported RLKs on *A. cepa* contigs. Contigs showing more than 40% homology with the bait RLKs were selected and the open reading frames (ORFs) and their putative protein sequences were obtained by using Expasy translate tool (<http://web.expasy.org/translate/>). Protein sequences obtained were used in a second round of BLAST search against the non-redundant protein database at NCBI to identify their closest homologs.

Sequence alignment and phylogenetic analysis

Multiple alignments of proteins deduced from the selected four contig sequences were performed using the Clustal Omega program (<https://www.ebi.ac.uk/Tools/msa/clustalo/>). Phylogenetic tree was constructed from the aligned contigs and other RLK protein sequences by employing the neighbor-joining method with Poisson correction, 1000 bootstrap replicates and pairwise deletion using the Molecular Evolutionary Genetic Analysis (MEGA v 7) package (Kumar et al., 2016).

Motif characterization, protein properties and hydrophobicity prediction

De novo motif predictions and motif elicitation of the selected contigs were performed by using Multiple Expectation Maximization for motif Elicitation (MEME) tool

(Bailey et al., 2006). Additional domains were detected using the Prosite (<http://prosite.expasy.org/>) and Pfam (<http://pfam.xfam.org/>) prediction programs. The peptide structural properties of the selected contigs were predicted by using Self Optimized Prediction Method with Alignment (SOPMA) (Combet et al., 2000). The hydrophobicity of the deduced protein sequences were predicted by using the Kyte-Doolittle plot (Kyte and Doolittle, 1982) (<http://web.expasy.org/protscale/>) keeping the window size 19.

Conclusion

The mining of *A. cepa* transcriptome database and extensive bioinformatics analysis resulted in identification of four highly homologous putative RLKs in onion. All four identified contigs possessed 9 out of 11 canonical kinase sub-domains. Prosite analysis results further revealed the presence of signature protein kinase domains and ATP-binding sites in all four contigs confirming them to be putative protein kinases. Phylogenetic analysis revealed that the contigs belong to different sub-families of RLKs, and their homology with previously known RLKs suggested their putative functionalities in *A. cepa*. Moreover, this work will serve as a framework to identify and characterize RLKs in onions involved in plant defense, as hardly any information is available on them. The full-length characterization and isolation of these putative RLKs will unearth their structural organization in *A. cepa*. Further biochemical and functional analysis of the characterized putative RLKs are necessary to confirm their involvement in stress and disease responses in onion.

Acknowledgment

The author is grateful to Dr. Raj Kumar Joshi for his encouragement and support. Thanks to Center for Biotechnology, Siksha 'O' Anusandhan University, Bhubaneswar for providing necessary research facilities and support.

Reference

- Afzal AJ, Wood AJ, Lightfoot DA (2008) Plant receptor-like serine threonine kinases: roles in signaling and plant defense. *Mol Plant-Microbe Interact.* 21(5):507-17.
- Amil-Ruiz F, Blanco-Portales R, Muñoz-Blanco J, Caballero JL (2011) The strawberry plant defense mechanism: a molecular review. *Plant Cell Physiol.* 52(11):1873-903.
- Andolfo G, Ercolano MR (2015) Plant innate immunity multicomponent model. *Front Plant Sci.* 6:978.
- Antolín-Llovera M, Petutsching EK, Ried MK, Lipka V, Nürnberger T, Robatzek S, Parniske M (2014) Knowing your friends and foes - plant receptor-like kinases as initiators of symbiosis or defence. *New Phytol.* 204(4):791-802.
- Bai L, Zhang G, Zhou Y, Zhang Z, Wang W, Du Y, Wu Z, Song CP (2009) Plasma membrane-associated proline-rich extensin-like receptor kinase 4, a novel regulator of Ca²⁺ signalling, is required for abscisic acid responses in *Arabidopsis thaliana*. *Plant J.* 60(2):314-27.
- Bailey TL, Williams N, Misleh C, Li WW (2006) MEME: discovering and analyzing DNA and protein sequence motifs. *Nucl Acids Res.* 34:W369-373.
- Baldwin S, Revanna R, Pither-Joyce M, Shaw M, Wright K, Thomson S, Moya L, Lee R, Macknight R, McCallum J (2014) Genetic analyses of bolting in bulb onion (*Allium cepa* L.). *Theor Appl Genet.* 127(3):535-47.
- Chand SK, Mohapatra RK, Joshi RK (2017) *In silico* identification and functional annotation of miRNAs and their targets from EST and GSS of onion (*Allium cepa* L.). *Can J Biotechnol.* 1:211-211.
- Chand SK, Nanda S, Rout E, Joshi RK (2015) Mining, characterization and validation of EST derived microsatellites from the transcriptome database of *Allium sativum* L. *Bioinformatics.* 11(3):145-50.
- Combet C, Blanchet C, Geourjon C, Deléage G (2000) NPS@: network protein sequence analysis. *Trends Biochem Sci.* 25(3):147-50.
- Emilia NS, Percy MC (2015) Gene mining a marama bean expressed sequence tags (ESTs) database: embryonic seed development genes and microsatellite marker identification. *African J Biotechnol.* 14(40):2836-2845.
- Food and agriculture organisation, statistics division (FAOSTAT) (2017) <http://www.fao.org/faostat>.
- Gish LA, Clark SE (2011) The RLK/Pelle family of kinases. *Plant J.* 66(1):117-27.
- Greeff C, Roux M, Mundy J, Petersen M (2012) Receptor-like kinase complexes in plant innate immunity. *Front Plant Sci.* 3:209.
- Guo H, Li L, Ye H, Yu X, Algreen A, Yin Y (2009) Three related receptor-like kinases are required for optimal cell elongation in *Arabidopsis thaliana*. *Proc Natl Acad Sci.* 106(18):7648-53.
- Guo H, Ye H, Li L, Yin Y (2009) A family of receptor-like kinases are regulated by BES1 and involved in plant growth in *Arabidopsis thaliana*. *Plant Signal Behav.* 4(8):784-6.
- Hanks SK, Quinn AM (1991) Protein kinase catalytic domain sequence database: identification of conserved features of primary structure and classification of family members. *Meth Enzymol.* 200:38-62.
- Ho HL (2015) Functional roles of plant protein kinases in signal transduction pathways during abiotic and biotic stress. *J Biodivers Biopros. Dev.* 2(2):147.
- Huang X, Madan A (1999) CAP3: A DNA sequence assembly program. *Genome Res.* 9(9):868-77.
- Jo J, Purushotham PM, Han K, Lee HR, Nah G, Kang BC (2017) Development of a genetic map for onion (*Allium cepa* L.) using reference-free genotyping-by-sequencing and SNP assays. *Front Plant Sci.* 8:1606.
- Joshi RK, Kar B, Nayak S (2011) Characterization of mitogen activated protein kinases (MAPKs) in the *Curcuma longa* expressed sequence tag database. *Bioinformatics.* 7(4):180-3.
- Joshi RK, Kar B, Nayak S (2011) Survey and characterization of NBS-LRR (R) genes in *Curcuma longa* transcriptome. *Bioinformatics.* 6(9):360-3.
- Kar B, Nayak S, Joshi RK (2012) Classification and comparative analysis of *Curcuma longa* L. expressed sequences tags (ESTs) encoding glycine-rich proteins (GRPs). *Bioinformatics* 8(3):142-6.

- Keeling CI, Weisshaar S, Ralph SG, Jancsik S, Hamberger B, Dullat HK, Bohlmann J (2011) Transcriptome mining, functional characterization, and phylogeny of a large terpene synthase gene family in spruce (*Picea* spp.). *BMC Plant Biol.* 11:43.
- Kulik A, Wawer I, Krzywińska E, Bucholc M, Dobrowolska G (2011) SnRK2 protein kinases—key regulators of plant response to abiotic stresses. *OMICS.* 15(12):859-72.
- Kumar S, Stecher G, Tamura K (2016) MEGA7: Molecular evolutionary genetics analysis version 7.0 for bigger datasets. *Mol Biol Evol.* 33(7):1870-4.
- Kyte J, Doolittle RF (1982) A simple method for displaying the hydrophobic character of a protein. *J Mol Biol.* 157:105–132.
- Lehti-Shiu MD, Shiu S-H (2012) Diversity, classification and function of the plant protein kinase superfamily. *Philos Trans R Soc B Biol Sci.* 367(1602):2619-39.
- Leng X, Han J, Wang X, Zhao M, Sun X, Wang C, Fang J (2015) Characterization of a calmodulin-binding transcription factor from strawberry (*Fragaria × ananassa*). *Plant Genome.* 8(2):1-12.
- Lin Z-JD, Liebrand TWH, Yadeta KA, Coaker GL (2015) PBL13 is a serine/threonine protein kinase that negatively regulates *Arabidopsis* immune responses. *Plant Physiol.* 169(4): 2950–2962.
- Lopez L, Barreiro R, Fischer M, Koch MA (2015) Mining microsatellite markers from public expressed sequence tags databases for the study of threatened plants. *BMC Genomics.* 16:781.
- Mengiste T (2012) Plant immunity to necrotrophs. *Annu Rev Phytopathol* 50:267–294.
- Nakhamchik A, Zhao Z, Provart NJ, Shiu SH, Keatley SK, Cameron RK, Goring DR (2004) A comprehensive expression analysis of the *Arabidopsis* proline-rich extensin-like receptor kinase gene family using bioinformatic and experimental approaches. *Plant Cell Physiol.* 45(12):1875-81.
- Nanda S, Nayak S, Joshi RK (2014) Molecular cloning and expression analysis of four turmeric MAP kinase genes in response to abiotic stresses and phytohormones. *Biol Plant.* 58(3):479–490.
- Nanda S, Rout E, Joshi RK (2016) *Curcuma longa* mitogen-activated protein kinase 6 (ClMPK6) stimulates the defense response pathway and enhances the resistance to necrotrophic fungal infection. *Plant Mol Biol Report.* 34(5):886–898.
- Ren G, Li X, Lin Z (2014) Mining, genetic mapping and expression analysis of EST-derived resistance gene homologs (RGHs) in cotton. *BMC Plant Biol.* 14:203.
- Rudrabhatla P, Reddy MM, Rajasekharan R (2006) Genome-wide analysis and experimentation of plant serine/threonine/tyrosine-specific protein kinases. *Plant Mol Biol.* 60(2):293-319.
- Sekhwil MK, Li P, Lam I, Wang X, Cloutier S, You FM (2015) Disease resistance gene analogs (RGAs) in plants. *Int J Mol Sci.* 16(8):19248-90.
- Shiu SH, Bleecker (2001) Plant receptor-like kinase gene family: diversity, function, and signaling. *Sci signal Transduct Knowl Environ.* 2001(113):1–13.
- Shiu S-H, Karlowski WM, Pan R, Tzeng Y-H, Mayer KFX, Li W-H (2004) Comparative analysis of the receptor-like kinase family in *Arabidopsis* and rice. *Plant Cell.* 16(5): 1220–1234.
- Silva NF, Goring DR (2002) The proline-rich, extensin-like receptor kinase-1 (PERK1) gene is rapidly induced by wounding. *Plant Mol Biol.* 50(4-5):667-85.
- Torto TA, Li S, Styer A, Huitema E, Testa A, Gow N, van West P, Kamoun S (2003) EST mining and functional expression assays identify extracellular effector proteins from the plant pathogen *Phytophthora*. *Genome Res.* 13(7):1675-85.
- Wako T, Yamashita K, Tsukazaki H, Ohara T, Kojima A, Yaguchi S, Shimazaki S, Midorikawa N, Sakai T, Yamauchi N, Shigyo M (2015) Screening and incorporation of rust resistance from *Allium cepa* into bunching onion (*Allium fistulosum*) via alien chromosome addition. *Genome.* 58(4):135-42.
- Xi Y, Pan PL, Ye YX, Yu B, Xu HJ, Zhang CX (2014) Chitinase-like gene family in the brown planthopper, *Nilaparvata lugens*. *Insect Mol Biol.* 24(1):29-40.