

THE RUBBER TREE KINOME: GENOME-WIDE CHARACTERIZATION AND INSIGHTS INTO CO-EXPRESSION PATTERNS ASSOCIATED WITH ABIOTIC STRESS RESPONSE

Lucas Borges dos Santos¹; Alexandre Hild Aono¹; Felipe Roberto Francisco¹; Livia Moura de Souza¹; Anete Pereira de Souza¹

¹ Center of Molecular Biology and Genetic Engineering (CBMEG), University of Campinas (UNICAMP), Campinas, São Paulo, Brazil

Conference: VII Congresso Brasileiro de Heveicultura -10 a 12 de novembro de 2021, Piracicaba/SP

Abstract: The protein kinase (PK) superfamily constitutes one of the largest and most conserved protein families in eucaryotic genomes, comprising core components of signaling pathways in cell regulation. Despite its remarkable relevance, only few kinase families have been studied in *Hevea brasiliensis* (Hb). In this study, we identified and characterized the entire set of PK genes, also known as kinome, present in the Hb genome. A total of 2,842 Hb PK genes were identified using a combination of two reference genome sources of the species. The kinome was further classified into 20 groups and 123 kinase subfamilies. We found that gene structure and protein properties of Hb PK subfamilies are highly comparable to the ones of two related species: *Manihot esculenta* and *Ricinus communis*. We identified 339 tandemly duplicated Hb PK genes and 362 PK genes associated with transposable elements. Quantification of the PK gene expression under a wide range of tissues and abiotic stress conditions (cold, drought, jasmonate and ethylene treatments) revealed remarkable differences in expression patterns of kinase subfamilies between latex and leaf tissues. In addition, co-expression networks modeled for control and abiotic-stressed samples revealed putative associations between several kinase subfamilies and abiotic-stress response.

Key-words: kinase, hevea brasiliensis, cold, drought.

Introduction

Different types of abiotic stresses may trigger several physiological responses in susceptible rubber tree genotypes and often impact its survival, growth and productivity, depending on the age and vigor of the affected plant (Kuruvilla et al., 2017). The ability to sense and adapt to adverse conditions rely on the activation of complex signaling networks that protect plants from potential damages caused by these environmental changes (Kovtun et al., 2000). Kinase proteins (PKs) comprise one of the most diverse protein superfamilies among eukaryotic organisms and participate as key components of the stimuli perception and signal transduction through a chain of phosphorylation events, resulting in the activation of genes and several cellular responses (Colcombet & Hirt, 2008). In rubber tree, several kinase families have been characterized, however, a deep characterization and global expression analysis of the PK superfamily is currently lacking.

In this study, we investigated the kinase diversity present in the *H. brasiliensis* (Hb) genome through a vast characterization of its kinase proteins. In addition to that, publicly RNA sequencing (RNA-Seq) data from 14 different Hb genotypes were used to identify expression patterns of the kinase subfamilies, followed by the construction of co-expression networks for control and abiotic stress conditions.

Material and Methods

We selected the two major genome assemblies of the rubber tree: the latest chromosome-level genome (Liu et al., 2020) and the reference scaffold-level assembly (Tang et al., 2016). Kinase proteins were identified based on similarities to the Hidden Markov Model of two kinase domains (PF00069 and PF07714). Hb kinome was created as a combination of putative kinase proteins identified from two different genomic datasets using CD-HIT. The same pipeline was further applied to *M. esculenta* (Me) and *R. communis* (Rc).

For each PK, we obtained the following characteristics: gene localization, intron number, molecular weight (Gasteiger et al., 2003), isoelectric point, subcellular localization, transmembrane domains, N-terminal signal peptides (Armenteros et al., 2019), and gene ontology terms (Götz et al., 2008). Tandem duplication events were determined based on the physical location of PK genes and their compositional similarities assessed through comparative alignments using the BLAST (Altschul et al., 1990). We also searched for transposable elements (TEs) in the Hb genome using BLAST.

Publicly available Hb RNA-Seq experiments were collected from the NCBI's Sequence Read Archive (SRA) database (Leinonen et al., 2010). PK expression was quantified using the Salmon software v.1.1.0 (Patro et al., 2017) and heatmaps for control and stress conditions were generated using the R package pheatmap (Kolde & Kolde, 2015). Co-expression networks were modeled and visualized using the R package igraph (Csardi & Nepusz, 2006) with a minimum Pearson correlation coefficient set to 0.7.

Results and Discussion

We identified 2,842 typical putative PK genes in Hb, 1,531 in Me, and 863 in Rc. These PKs were further classified into groups and subfamilies, based on HMM profiles of 127 kinase subfamilies defined by Lehti-Shiu & Shiu (2009), and further validated through phylogenetic analyses (Figure 1). The RLK-Pelle is the most represented group under all three species, divided into 59 different subfamilies and which accounts for 65.5%, 68.1%, 65.2% of all Hb, Me, and Rc PKs, respectively. Most interestingly, the number of kinase-encoding genes found in Hb is comparable to the ones found in species with larger and polyploid genomes, including sugarcane (2,919 PKs; Aono et al., 2021), wheat (3,269 PKs; Yan et al., 2017) and cotton (1,517 to 2,745 PKs; Chen et al., 2018), which suggests that kinome complexity is not limited to genome size and ploidy, but also affected by other evolutionary processes such as recent whole-genome duplications.

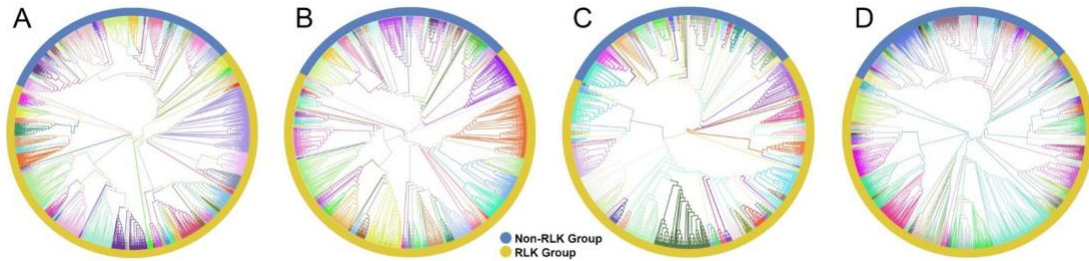


Figure 1. (A) Phylogenetic tree constructed with 2,842 Hbr PKs organized in 123 subfamilies. (B) Phylogenetic tree of the 1,531 Me PKs organized in 123 subfamilies. (C) Phylogenetic tree of the 863 Rc PKs organized in 125 subfamilies. (D) Phylogenetic tree of all Hb, Me, and Rc PKs. Kinase subfamilies are represented by different branch colors.

Protein characteristics of all three kinomes were highly comparable. Several PKs have predicted transmembrane domains (45.6% in Hb, 50.5% in Me, and 48.2% in Rc), N-terminal signal peptides (29.6%, 37.3%, and 33.5%, respectively). Similarly, the distribution of molecular weights and isoelectric points was relatively even (Figure 2). However, it is worth mentioning that Me is more similar to Hb in kinome size than Rc. An increase in the kinome size could be partly attributed to the expansion of several gene families throughout duplication events during their evolutionary history (Lehti-Shiu & Shiu, 2012).

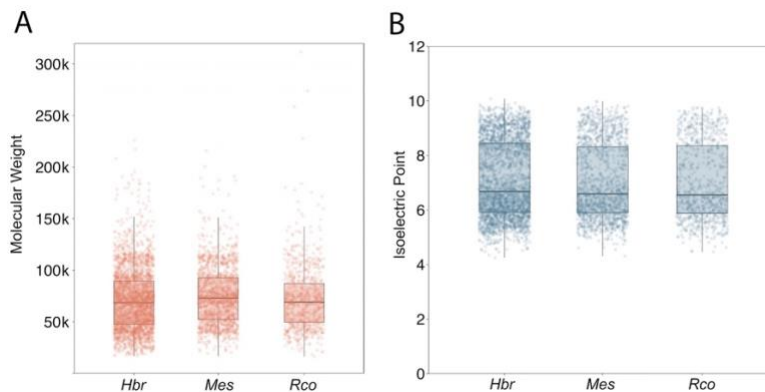


Figure 2. (A) Molecular weight (MW) and (B) Isoelectric point (pI) distribution of *H. brasiliensis* (Hbr), *M. esculenta* (Mes), and *R. communis* (Rco) PKs.

The quantification analysis revealed different expression profiles of PK subfamilies among different tissues, genotypes and conditions. PK subfamilies STE_STE-PI and RLK-Pelle_LRR-VIII-1 showed increased expression in bark and latex when compared to leaf. Moreover, the RLK-Pelle_WAK_LRK10L-1 subfamily showed an expression increase in leaf after the exposure to cold stress; and transcript levels of Group-PI-4 increased in response to drought and low temperatures. Ultimately, co-expression networks of control and abiotic-stress conditions revealed different patterns of interaction between PKs during control and abiotic-stress situations (Figure 3). We found differences in the most influential subfamilies within each network, where during conditions of abiotic-stress, PK subfamilies of the RLK-Pelle group (RLK-Pelle_RLCK-X, RLK-Pelle_PERK-1, RLK-Pelle_LysM, RLK-Pelle_SD-2b, RLK-Pelle_L-LEC, RLK-Pelle_RLCK-VIIa-1), CAMK_CDPK, TKL-PI-4, STE_STE7, and CK1-CK1-PI have the largest number of expression correlation (edges) within the network, thus suggesting an important participation during the process of physiological adaptation during adverse conditions.

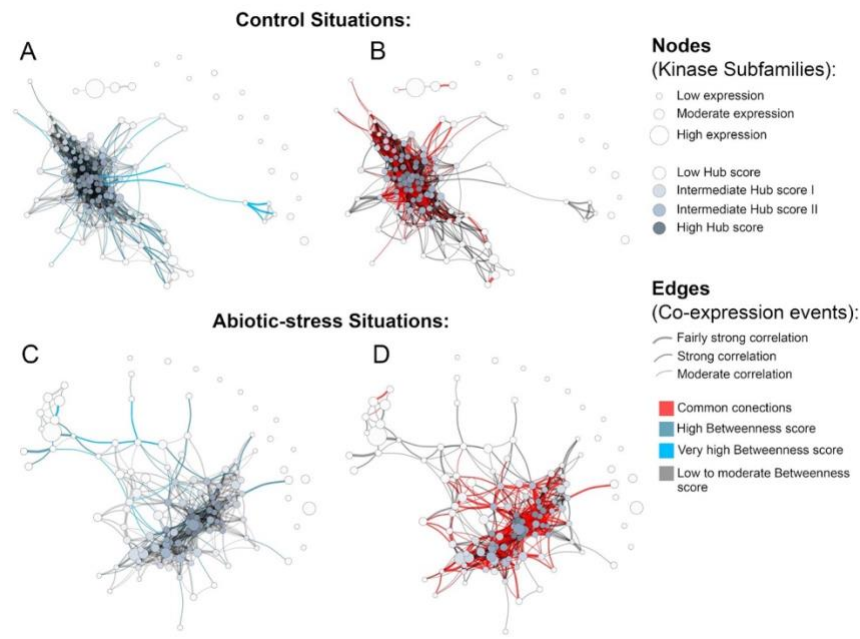


Figure 3: Co-expression networks for *H. brasiliensis* (Hb) kinase subfamilies. (A) Hb control network with betweenness values highlighted in blue. (B) Hb control network indicating edge similarities (red) with the Hb stress network. (C) Hb abiotic-stress network with betweenness values highlighted in blue. (D) Hb abiotic-stress network indicating edge similarities (red) with the Hb control network.

Conclusion

These results provide a large reservoir of data for rubber tree research, together with new insights of key components involved during plant response to abiotic factors that are responsible for significant reductions in latex yield. While most PK subfamilies showed relatively similar patterns of expression during regular and abiotic-stress situations, some PK subfamilies could be potentially associated with stress perception and the activation of several biological pathways that result in physiological changes in rubber tree.

Acknowledgements

We acknowledge The São Paulo Research Foundation (FAPESP) and the Coordination for the Improvement of Higher Education Personnel (CAPES) for funding this research project.

References

- ALMAGRO ARMENTEROS, JOSÉ JUAN, TSIRIGOS, KONSTANTINOS D.SØNDERBY, CASPER KAAE et al. SignalP 5.0 improves signal peptide predictions using deep neural networks. **Nature Biotechnology**, v. 37, n. 4, p. 420-423, 2019.
- ALTSCHUL, STEPHEN F., GISH, WARRENMILLER, WEBB et al. Basic local alignment search tool. **Journal of Molecular Biology**, v. 215, n. 3, p. 403-410, 1990.
- AONO, ALEXANDRE HILD, PIMENTA, RICARDO JOSÉ GONZAGAGARCIA, ANA LETYCIA BASSO et al. The Wild Sugarcane and Sorghum Kinomes: Insights Into Expansion, Diversification, and Expression Patterns. **Frontiers in Plant Science**, v. 12, 2021.
- CHEN, MINGYANG, LI, KUNLI, HAIPENG et al. The Glutathione Peroxidase Gene Family in *Gossypium hirsutum*: Genome-Wide Identification, Classification, Gene Expression and Functional Analysis. **Scientific Reports**, v. 7, n. 1, 2017.
- COLCOMBET, JEANHIRT, HERIBERT. Arabidopsis MAPKs: a complex signalling network involved in multiple biological processes. **Biochemical Journal**, v. 413, n. 2, p. 217-226, 2008.
- CSARDI, GABOR, NEPUSZ, TAMAS. The igraph software package for complex network research. **InterJournal, complex systems**, v. 1695, n.5, p.1-9, 2006.
- GASTEIGER, E. ExPASy: the proteomics server for in-depth protein knowledge and analysis. **Nucleic Acids Research**, v. 31, n. 13, p. 3784-3788, 2003.

- GOTZ, S., GARCIA-GOMEZ, J. M. TEROL, J. et al. High-throughput functional annotation and data mining with the Blast2GO suite. **Nucleic Acids Research**, v. 36, n. 10, p. 3420-3435, 2008.
- KOLDE, RAIVO, MAINTAINER, RAIVO KOLDE. Package ‘pheatmap’. **R Package**, v. 1, n. 7, p. 790, 2015.
- KOVTUN, Y., CHIU, W.-L. TENA, G. et al. Functional analysis of oxidative stress-activated mitogen-activated protein kinase cascade in plants. **Proceedings of the National Academy of Sciences**, v. 97, n. 6, p. 2940-2945, 2000.
- KURUVILLA, LINU, SATHIK, MB MOHAMED THOMAS, MOLLY et al. Identification and validation of cold responsive microRNAs of *Hevea brasiliensis* using high throughput sequencing. **Journal of Crop Science and Biotechnology**, v. 20, n. 5, p. 369-377, 2017.
- LEHTI-SHIU, MELISSA D., ZOU, CHENGHANADA, KOUSUKE et al. Evolutionary History and Stress Regulation of Plant Receptor-Like Kinase/Pelle Genes. **Plant Physiology**, v. 150, n. 1, p. 12-26, 2009.
- LIU, JIN, SHI, CONGSHI, CHENG-CHENG et al. The Chromosome-Based Rubber Tree Genome Provides New Insights into Spurge Genome Evolution and Rubber Biosynthesis. **Molecular Plant**, v. 13, n. 2, p. 336-350, 2020.
- PATRO, ROB, DUGGAL, GEETLOVE, MICHAEL I et al. Salmon provides fast and bias-aware quantification of transcript expression. **Nature Methods**, v. 14, n. 4, p. 417-419, 2017.
- TANG, CHAORONG, YANG, MENGFANG, YONGJUN et al. The rubber tree genome reveals new insights into rubber production and species adaptation. **Nature Plants**, v. 2, n. 6, 2016.
- YAN, JUN, SU, PEISENWEI, ZHAORAN et al. Genome-wide identification, classification, evolutionary analysis and gene expression patterns of the protein kinase gene family in wheat and *Aegilops tauschii*. **Plant Molecular Biology**, v. 95, n. 3, p. 227-242, 2017.