Thesis of the PhD dissertation

Ribal Masri

Gödöllő

2023

Hungarian University of Agriculture and Life Sciences

A transcriptomic study exploring the role of the *Vitis vinifera NAC* genes in response to stresses: case of *Botrytis cinerea*

Thesis of the PhD dissertation

Ribal Masri

Gödöllő

2023

The PhD School

Name: Doctoral School of Biological Sciences

Discipline: Plant Biotechnology

Head: **Prof. Dr. Zoltán Nagy**
Director of the institute, professor
MATE, Gödöllő

Supervisor: **Prof. Dr. Erzsébet Kiss**
Professor emeritus
MATE, Gödöllő
Institute of Genetics and Biotechnology

....	....
Prof. Dr. Zoltán Nagy	**Prof. Dr. Erzsébet Kiss**
Approval of the School Leader	Approval of the Supervisor

1. INTRODUCTION

Biotic and abiotic stresses are the main factors that limit growth and fruit quality of a significant proportion of essential crops worldwide. Plants have developed the ability to cope with the harmful effects of stresses by activating a number of defense mechanisms which include phytohormone signaling networks (Masri and Kiss 2023).

Transcription factor (TFs) proteins are essential regulators of gene expression in all living organisms, with roles in plant growth, cell cycling, cell signaling, and stress response (Li et al., 2021) by binding to specific short sequence motifs mostly detected within the promoters of the target genes, known as cis-regulatory sequences (Priest et al., 2009). Thus, the interaction between them, formed a crucial functional link of gene regulatory networks in plant defence response by controlling various stresses (Kaur et al., 2016).

The NACs (**N**AM – no apical meristem, **A**TAF – *Arabidopsis* transcription activation factor, and **C**UC – cup-shaped cotyledon) are considered as one of the largest TF families that consists of N-terminus and C-terminal. The NAC domain is highly conserved, comprise of 160 amino acid residues that classified into five subdomains from A to E (Puranik et al., 2012). The researchers discovered several *NAC* genes from various plant species using genomic sequencing, including 151 in rice, 117 in *Arabidopsis*, 74 in grape (Shao et al., 2015; Ju et al., 2020).

Different roles for NAC proteins in response to abiotic and biotic stresses, as well as in developmental processes have been reported such as, leaf senescence (Breeze et al. 2011), cell division (Kim et al. 2006), seed development (Sperotto et al. 2009), fiber development (Ko et al. 2007), shoot apical meristem formation (Kim et al. 2006) and embryo development (Duval et al. 2002).

Grapevine (*Vitis vinifera* L.), a member of the *Vitaceae* family and one of the world's most valuable fruit crops with high economic value is proned to a wide range of pathogens that cause yield and quality losses (Ju et al., 2020).

Plant pathogens are categorized to different classes according to their modes of nutrition, such as: biotrophic, necrotrophic, hemibiotrophic and saprophytic pathogens. The necrotrophic fungus *Botrytis cinerea,* is the main cause of the grey mold disease in more than 200 plants (Elad et al., 2004) and one of the most crucial pathogens in grapevine (Haile et al., 2017) with a short biotrophic phase (Veloso and van Kan, 2018).

However, in the vineyards, the infection of *B. cinerea* targets primarily the ripe berries; the pathogen inoculation occurs during the onset of the grape berry development (Keller et al., 2003; Pezet et al., 2003).

The *B. cinerea* weakens the natural pathogen resistance system of grape berries and imposes different modifications such as lowering the mechanical resistance of the cell-wall, compacting the bunches and increasing the sugar levels and the concentration of organic acids (Prusky et al., 2013; Blanco-Ulate et al., 2015). Additionally, speeding up the ripening process or promoting the programmed cell death machinery of the grape berries can be manipulated by this pathogen during the infection (Veloso and van Kan, 2018).

Many studies have reported the role of *V. vinifera NAC* genes in response to biotic stress. For example, the *VvNAC1* gene *(*known as *VvNAC60* in Genoscope database*)* showed resistance to *B.*

cinerea and *Hyaloperonospora arabidopsidis* (Le Henanff et al., 2013), *VvNAC36* gene in *A. thaliana* was up-regulated in response to powdery mildew colonization (Tóth et al., 2016).

Therefore, studying NAC transcription factor promoters can reveal useful information about the genes and signaling networks involved in abiotic and biotic stress responses. Web-based databases such as, Genomatix, PLACE and Plant CARE have provided a convenient and easy *in silico* way to search motifs and detect cis-acting regulatory elements (CAREs) of the promoters (Ibraheem et al., 2010). Recently, many studies for elucidation and annotation of gene functions have been carried out by use of *in silico* methods. Fortunately, the past decade has seen a revolution in omics technologies that have generated abundant amounts of useful data for *in silico* function predictions (Rhee et al., 2014).

2. OBJECTIVES

The NAC proteins play an important role in providing resistance to plants. Till now, few works have been reported on cis-acting regulatory elements present in *Vitis NAC* genes. Therefore, our aim in the present work was to perform a transcriptomic study of the *VvNAC* genes to report their responses against *B.cinerea* through an *in silico* approach by characterizing the cis-acting regulatory elements of the *Vitis vinifera* L. *NAC* genes, followed by an *in vitro* work wherever is available in order to find a link between both approaches giving results that might be complemented. This work throws light on the promoter regions of grape *NAC* genes which further can provide new ways for the plant genetic engineering technology for protection of crops against stresses and especially *B.cinerea.*

2.1. Characterizing of the *Vitis vinifera* L. *NAC* TF genes through an *in silico* approach.

- Identifying the *c*is-acting regulatory elements (CAREs) of the *Vitis vinifera* L. *NAC* genes.
- Analyzing the microarray expression profiles of the *Vitis vinifera* L. *NAC* genes in response to *B. cinerea*.
- Interpreting the up-regulations results by linking them to the existence of specific CAREs in the promoters of the up-regulated *VvNAC* genes.

2.2. Detailed interpretation of the function of the *VvNAC* genes.

- Detecting the remarkable responses whereas up-regulation or down- regulation of the *VvNAC* genes is induced by *B. cinerea.*
- Predicting new genes with a function that might respond to stresses and especially to *B. cinerea*.

2.3. Generating *in vitro* approach.

- Determination of the subcellular localization through a transient expression for the *VvNAC36* gene in a model tobacco plant, *Nicotiana benthamiana* L.
- Comparing the transient expression results with the computational results to validate or find a link between both approaches.

3. MATERIALS AND METHODS

3.1. *In silico* part

3.1.1. Detection and data analyses of NAC family genes

A local blast analysis was conducted to locate *NAC* genes within the grapevine genome. The NAM domain and 26,346 predicted genes from the 12X assembled *Vitis vinifera* Pinot Noir PN40024 genome were utilized as search queries and databases. As a result, 74 genes were successfully identified and designated from *VvNAC01* to *VvNAC74*, each associated with a unique gene locus number (Wang et al., 2013).

In this study, we utilized as input the gene loci of the 74 *Vitis vinifera NAC* genes to retrieve the protein sequences from genoscope, the grapevine genomic database available at (http://www.genoscope.cns.fr/externe/GenomeBrowser/Vitis/).

The NAC conserved domain sequences of all the genes were extracted through the Pfam web database (https://pfam.xfam.org) (Finn et al., 2016). Subsequently, a BLASTp analysis was performed using the phytozome v13 database accessible at https://phytozome-next.jgi.doe.gov/ to validate all VvNAC protein sequences employed in this approach (Goodstein et al., 2012).

The amino acid composition, isoelectric point, and molecular weight of the NAC proteins were determined using ExPASy; http://ca.expasy.org/tools/pi_tool.html (Ison et al., 2013). CELLO v.2.5 webtool; http://cello.life.nctu.edu.tw was utilized to estimate the subcellular localization of the proteins (Yu et al., 2004).

TMHMM v2.0 (http://www.cbs.dtu.dk/services/TMHMM/) was used to determine the transmembrane helices (TMHs) of the identified *VvNAC* family proteins (Krzywinski et al., 2009).

TMHs are regions within a protein that span the lipid bilayer of a cell membrane. By analyzing the amino acid sequence of the *VvNAC* genes, TMHMM v2.0 uses computational algorithms and statistical models to identify potential TMHs within the protein. The prediction is based on specific features such as hydrophobicity and amphipathicity, which are characteristic of transmembrane regions.

The nucleotide substitution rates (Ka and Ks) and their ratios (Ka/Ks) for duplicated *VvNAC* genes were computed using the straightforward Ka/Ks calculator feature within the TBtools software, which can be found at https://bio.tools/tbtools (Chen et al., 2020). To facilitate this analysis, the coding sequence (CDS) sequences of the genes were sourced from the phytozome v13 database.

3.1.2. Conserved motifs and cis-acting regulatory elements (CAREs) of VvNAC genes

The protein motifs of *Vitis vinifera NAC* were analyzed using the web database program MEME 5.5.2, which can be accessed at https://meme-suite.org/meme/tools/meme. This program utilizes various statistical modeling techniques to identify prevalent patterns, 10 Motifs were selected based on the criteria of a minimum width of 6 pixels and a maximum width of 50 pixels (Bailey et al., 2009).

The cis-acting regulatory elements of the *VvNAC* genes were identified using the Plant CARE database, which is accessible at http://bioinformatics.psb.ugent.be/webtools/plantcare/html/ (Lescot et al., 2002).

3.1.3. Protein–protein interaction analysis

In this study, we utilized the web-based STRING (Search Tool for the Retrieval of Interacting Genes/Proteins) database version 11.0 https://string-db.org to predict and construct protein-protein interaction networks involving VvNAC proteins, with *Arabidopsis* homologous proteins as references (Szklarczyk et al., 2023).

To configure the STRING analysis, we adjusted specific parameters as follows: the network type was defined as the full STRING network, the meaning of network edges was set to evidence, the minimum required interaction score was established as a medium confidence parameter (0.4), and the maximum number of interactions displayed for each protein did not exceed 10 interactors.

3.1.4. In silico expression profiles of VvNAC genes and statistical analysis

The microarray dataset of grapevine *NAC* gene expression in response to *Botrytis cinerea* stress was obtained from the open science resource Expression Atlas, which can be accessed at https://www.ebi.ac.uk/gxa/experiments/E-GEOD-67932/Results (Blanco-Ulate et al., 2015) (Appendix - Table 7.4). Expression patterns of these *NAC* genes were represented as a heat map viewed in TBtools https://bio.tools/tbtools.

Additionally, by using Bioinformatics.com.cn a statistical analysis for gene expression through hierarchical clustering data was generated.

3.2. *In vitro* part

3.2.1. Construction of plant expression vectors

VvNAC36 gene construct

Based on previous work started in the institute, the *VvNAC36* has been chosen to check how much accurate the *in silico* results might be to compare to the laboratory work. Thus, genomic DNA was extracted from young grapevine leaves using the CTAB (Cetyltrimethylammonium bromide) method by following these steps:

1. Sample collection: We begin by collecting plant material, such as leaves, young shoots, or other tissues of interest, ensuring that the material is fresh and free from contamination.

2. Homogenization: We grind the collected plant material into a fine powder, using either a mortar and pestle or specialized homogenization equipment. Our goal is to break open the cells and release their contents.

3. Cell lysis: We transfer the powdered plant material into a tube or container and add a CTAB-based extraction buffer. The extraction buffer typically contains CTAB, EDTA (Ethylene Diamine Tetraacetic Acid), and other reagents to disrupt cell membranes and stabilize DNA. We then heat the mixture to facilitate cell lysis.

4. Incubation: After adding the extraction buffer, we incubate the mixture at an elevated temperature, usually around 65°C, for a specified period. This step helps break down cell membranes and release DNA into the solution

5. Phenol-Chloroform extraction: Following incubation, we add an equal volume of phenol-chloroform-isoamyl alcohol (PCI) to the mixture. Phenol-chloroform helps to separate DNA from proteins, lipids, and other cellular components. We mix the solution thoroughly by inverting the tube or using a vortex mixer.

6. Centrifugation: We then centrifuge the mixture at high speed. This separates it into three phases: a lower organic phase (containing proteins and lipids), an interphase, and an upper aqueous phase (containing DNA). Carefully, we transfer the upper aqueous phase to a new tube.

7. Precipitation: To precipitate DNA, we add cold isopropanol or ethanol to the aqueous phase. DNA is not soluble in these alcohols, causing it to precipitate out of solution. We gently invert the tube to encourage DNA precipitation.

8. Centrifugation: After precipitation, we centrifuge the tube again to pellet the DNA. We then remove the alcohol-containing supernatant, taking care not to disturb the DNA pellet.

9. Washing: We wash the DNA pellet with cold ethanol to remove any residual contaminants, which further purifies the DNA

10. Drying: We allow the DNA pellet to air dry for a brief period, ensuring it's completely dry to avoid ethanol contamination in downstream applications

11. Rehydration: To dissolve the DNA, we rehydrate the DNA pellet in an appropriate buffer, such as TE buffer (Tris-EDTA). This buffer helps protect the DNA from degradation.

12. Storage: Finally, we store the extracted DNA at -20°C or -80°C for long-term preservation or at 4°C for short-term use.

Moreover, the *VvNAC36* gene was then amplified by PCR using Phusion High Fidelity Taq polymerase with forward and reverse primers, which contained *HindIII* and *SacI* restriction sites, respectively.

The PCR amplification was performed for 34 cycles under the following conditions: 98°C for 30 s, 94°C for 10 s, 60°C for 90 s, and 72°C for 5 min, followed by a final extension step at 72°C for 10 min. The purified PCR product was cloned into the pENTR/D-TOPO vector from Invitrogen, further digested, and subcloned at the *HindIII* and *SacI* restriction sites into the pGWB605 binary vector which contains a CaMV35S promoter and the synthetic green fluorescent protein (sGFP) reporter gene, resulting in the recombinant plasmid pGWB605::*VvNAC36*.

The control vector 35S-sGFP (containing green fluorescent protein alone) and the recombinant plasmid pGWB605::*VvNAC36* were introduced into *Agrobacterium tumefaciens* strain GV3101 by direct transformation.

3.2.2. Transient Expression

For the pre-culture, *Agrobacterium tumefaciens* (*A. tumefaciens*) GV3101 carrying the binary vector pGWB605::*VvNAC36* and 35S-sGFP were grown overnight in 2–3 mL Luria-Bertani (LB medium) supplemented with 100 mg/L spectinomycin and 30 mg/L rifampicin at 28 °C (Gusain et al., 2021) . Subsequently, 1 mL of the pre-cultures was inoculated into fresh 25 mL LB medium containing the same antibiotics for the main cultures.

On the following day, culture solutions were centrifuged at 5,000 rpm for 15 minutes at room temperature, followed by resuspension of the *A. tumefaciens* pellet in a buffer consisting of 1 M

$MgCl_2$, 0.5 M 2-(N-morpholino)ethanesulfonic acid (MES) (pH 5.6), and 10 mM acetosyringone. The buffer volume was adjusted to achieve an OD 600 of approximately 1 (Hoshikawa et al., 2018). The buffer was then incubated at room temperature for 2–3 hours.

Using a needleless 1-mL syringe, the cells were co-infiltrated at a 1:1 ratio with *A. tumefaciens* containing the p19 gene, a silencing suppressor (Lindbo, 2007), into the abaxial side of the leaves of 6-week-old plants.

Following infiltration, the plants were left undisturbed for 5 days before examination with the Leica TCS SP8 confocal microscope. This examination involved analyzing pieces of infiltrated leaves taken from the infected region at the National Agricultural Research and Innovation Centre (NARIC) in Hungary.

4. RESULTS

4.1. Identification and genomic distribution of NAC proteins

In this study, different databases were utilized to study and characterize the *Vitis vinifera's NAC* proteins. The genome of *V. vinifera* L. was found to have 74 NAC proteins.

The results presented a comprehensive summary of the *VvNAC* genes, including their length, molecular weight (MW), isoelectric point (pI), and cellular location. The length of the proteins produced by *NAC* genes ranged from 128 amino acids (AA) in *VvNAC50* to 994 (AA) in *VvNAC54*. The molecular weight of these proteins varied from 14.76 to 110.67 kDa (*VvNAC50* and *VvNAC54*, respectively). The isoelectric point of the *NAC* proteins ranged from 4.1 in *VvNAC47* to 10.52 in *VvNAC73*.

4.2. TMHs and KaKs value of *VvNAC* genes

Notably, nine out of the 74 *VvNAC* genes identified encoded proteins with transmembrane helices (TMHs). It was found that eight proteins, namely *VvNAC15, VvNAC20, VvNAC25, VvNAC29, VvNAC48, VvNAC54, VvNAC67,* and *VvNAC71* contained only one transmembrane domain while *VvNAC72* contained 6 domains.

Therefore, this information is valuable because the presence and location of transmembrane helices can provide insights into the protein's function and its association with cellular membranes. It suggests that these *VvNAC* proteins are likely to be membrane-associated or have specific interactions with the cell membrane.

Based on the analysis conducted using the KaKs calculator via TBtools, none of the 20 tandem repeat sequences exhibited Ka/Ks values greater than 1. This outcome implies the absence of positive selection in any of the tandem repeat sequences, which are typically associated with facilitating adaptive genetic variation and potentially influencing species evolution.

Instead, all genes displayed Ka/Ks values below 1, indicating the prevalence of negative selection during the course of evolution. This negative selection process serves to reduce the rate of changes in the amino acid profile. In summary, the findings indicate that the majority of *VvNAC* genes exhibit a slow evolutionary rate, characterized by the absence of positive selection.

4.3. Conserved motifs of *VvNAC* genes

The *VvNAC* genes contain a highly conserved DNA binding NAC domain at the N-terminal, consisting of 160 amino acid residues that are divided into five subdomains: A, B, C, D, and E. However, the C-terminal part is highly variable and lacks any known protein domains.

To further investigate the conserved motifs of NAC proteins from *Vitis vinifera* L., the MEME 5.5.2 web database program was utilized, resulting in the identification of one to ten motifs. A thorough investigation of the conserved motifs present in each *VvNAC* gene was also performed.

The outcomes of the investigation pertaining to the distribution of NAC protein motifs have elucidated significant patterns. Specifically, within the subset of 74 grape NAC proteins, a substantial portion comprising 57 proteins (equivalent to 77% of the total) has been identified as encompassing subdomains A to E. Furthermore, a distinct profile has emerged where six proteins exhibit the absence of a singular NAC sub-domain, namely A, B, or C. In addition, an analogous group of six proteins is characterized by the lack of two sub-domains, manifesting either as B and C or A and C.

A remarkable observation pertains to the presence of two sub-domains, A and B, within a sole protein, *VvNAC66.* Moreover, a noteworthy subset consists of three proteins, specifically *VvNAC35, VvNAC38*, and *VvNAC50*, each harboring only a solitary sub-domain, namely the C sub-domain. Notably, another protein, *VvNAC72*, distinctly carries the A sub-domain.

Moreover, the range of motifs identified within *Vitis* NAC proteins extended from one to thirteen. Remarkably, six VvNAC proteins namely, VvNAC28, VvNAC30, VvNAC31, VvNAC32, VvNAC36, and VvNAC60 hosted eight motifs each. In parallel, four proteins, specifically VvNAC09, VvNAC35, VvNAC38, and VvNAC66, showcased the presence of three motifs, while VvNAC50 featured two motifs.

Notably, motifs 1 and 7 exhibited a consistent occurrence among the VvNAC proteins. The analysis unveiled six motifs prevalent in 52.7% of the 74 VvNAC proteins, whereas seven motifs were identified in 9.4%, and five motifs presented in 20.2% of the VvNAC proteins. It's noteworthy to mention that VvNAC54 was notable for harboring the highest count of motifs, specifically thirteen, while conversely, VvNAC72 exhibited the lowest, featuring just a solitary motif.

4.4. Protein–protein interaction of *VvNAC*

The predicted protein-protein interaction map revealed intricate interactions among proteins themselves as well as with several other proteins. To identify optimal matches, NAC protein sequences from *Vitis vinifera* L. were compared against *Arabidopsis thaliana* proteins, and the corresponding proteins were selected for subsequent network analysis.

The VvNACs protein interaction network was established based on *Arabidopsis* protein orthologs, and those VvNAC proteins exhibiting high similarity to *Arabidopsis* proteins were referred to as STRING proteins.

Specifically, VvNAC26 and VvNAC43 displayed homology with NAP, VvNAC67 with NTL8, VvNAC51, VvNAC52, VvNAC53, VvNAC54, VvNAC55, VvNAC69, and VvNAC71 with

NTL9, VvNAC64 with NAC028, and VvNAC60 with NAC047. Notably, all of these proteins exhibited significant interactions with the NTL1 protein transporter activity, highlighting their functional relevance.

Furthermore, within a distinct cluster, VvNAC47 and VvNAC58 showed homology with XND1, VvNAC63 with VND7, VvNAC24 and VvNAC49 with NST1, VvNAC01, VvNAC07, and VvNAC73 with NAC083, and VvNAC34 and VvNAC37 with NAC073. These proteins displayed robust associations with various regulatory factors, including MYB46 (involved in regulating secondary wall biosynthesis in fibres and vessels), MYB83 (a key player in the NAC012/SND1-mediated transcriptional network governing secondary wall biosynthesis), MYB80 (critical for the timing of tapetal programmed cell death, a pivotal process in pollen development), MYB85 (a putative transcription factor), AtMYB103 (a putative MYB family transcription factor), and KNAT7 (a KNOTTED-like homeobox protein in *Arabidopsis thaliana* 7, implicated in secondary cell wall biosynthesis).

Moreover, VvNAC21 and VvNAC50 exhibited homology with NAC050, a DNA-binding transcription factor involved in multicellular organismal development and transcription regulation. These proteins demonstrated strong associations with JMJ14, a transcriptional repressor involved in DRM2-mediated maintenance of DNA methylation, as well as with AT1G72460, a leucine-rich repeat protein kinase family protein participating in the transmembrane receptor protein tyrosine kinase signalling pathway and protein amino acid phosphorylation.

Finally, VvNAC08 and VvNAC39, homologous to ATAF1, acted as pivotal genes connecting the two clusters and exhibited strong interactions with VvNAC44 (homologous to NAC032), which positively regulates age-dependent senescence, dark-induced leaf senescence, and stress-induced senescence.

Consequently, based on STRING database, the Gene Ontology (GO) predictions revealed various gene functions within the network, including the nucleus, DNA binding, organic cyclic compound binding, regulation of cellular metabolic processes, regulation of macromolecule metabolic processes, regulation of transcription (DNA-templated), regulation of nucleobase-containing compound metabolism, cell cycle, and DNA repair.

4.5. Promoter regions detection and analysis of cis-acting regulatory elements

To identify cis-acting regulatory elements (CAREs) in *Vitis vinifera's NAC* genes, the promoter sequences were analyzed using the Plant Care database. Specifically, the sequences were examined up to 1500 base pairs (bp) upstream of the translation start site. CAREs are small, conserved patterns of nucleotides that are crucial components of genes, commonly found at the 5' ends.

In total, 74 *Vitis vinifera* CAREs were analyzed to explore their response to diverse stress types. The investigation revealed that the *Vitis vinifera NAC* genes contain more than 3000 CAREs. The identified CAREs include crucial elements that respond to both biotic and abiotic stressors, involve in hormonal regulation, assist in the development, and perform various functions in the promoter regions of *VvNAC* genes.

4.5.1. Stress responsive cis-acting regulatory elements

To begin, abiotic stress-related motifs, such as MYC, MYB, and MBS-MYB binding sites, LTR (low-temperature elements), STRE (stress-responsive elements), and DRE (dehydration-responsive elements), were identified. Additionally, various CAREs were discovered in response to biotic stress, including STRE, wound-responsive (WUN motif and WRE3), elicitor-responsive (W box), and TC-rich repeats (defense and stress).

Furthermore, the study demonstrated that all *Vitis NAC* promoters contained motifs specific to light stresses, such as the Sp1, AE-box, Box 4, Chs-CMA1a, GA motif, GT1 motif, GATA-motif, I-box, Box S, G box, Gap box, and Box II.

4.5.2. Cis-acting regulatory elements in hormonal regulation.

Moving on, the cis-regulatory elements of the *VvNACs* were found to be involved in hormonal regulation, including multiple ligand-responsive elements such as the TGA-box (auxin responsive elements), ERE (ethylene responsive elements), GARE motifs (gibberellin responsive elements), P-box, ABRE (abscisic acid-responsive elements), AuxRR-core, TCA elements (salicylic acid-responsive elements), TATC-box, and TGA-element. These elements enhance the plants' stress resistance capabilities.

4.5.3. Role of cis-acting regulatory elements in cellular development:

Finally, while the number of motifs involved in cellular development is relatively fewer compared to other CAREs, the presence of several elements has been demonstrated. These include the CAT box (for meristem-specific activation), Motif I, GCN4 motif (for endosperm expression), AC-I and AC-II (for xylem-specific expression), O2-site (for zein metabolism regulation), F-box (for reproductive and vegetative development and growth, as well as cell defense and death), ARE (for anaerobic induction), the circadian motif, HD-Zip1 (for leaf development), AP-1 (for inducible flowering), and MBSI (for flavonoid biosynthesis gene regulation).

Given that most *VvNACs* exhibit three distinct types of stress-responsive CAREs, it is plausible to assume that these genes play a crucial role in enhancing tolerance to various biotic and abiotic stresses. Overall, these findings suggest that *Vitis vinifera's NAC* genes contain a diverse range of CAREs that respond to various stressors and are involved in hormonal regulation and cellular development.

4.5.4. In silico analysis of VvNAC gene expression

<u>Biotic expression</u>

The transcription profiling by high throughput sequencing of grapevine *NAC* genes at 3 different stages (early S1, middle S2, and late S3) of noble rot caused by *Botrytis cinerea* was extracted from the Expression Atlas database.

The analysis of the microarray data set showed that several *VvNAC* genes, including *VvNAC39*, *VvNAC44* and *VvNAC73* were highly regulated during stages S2 and S3. Additionally, there was a slight up-regulation in *VvNAC08*, *VvNAC30, VvNAC31, VvNAC36, VvNAC41, VvNAC60, VvNAC61* and *VvNAC74.* Conversely, there was a significant down-regulation in *VvNAC13, VvNAC51, VvNAC52, VvNAC53* and *VvNAC54*.

These genes were found to be enriched with biotic stress-related motifs, thus highlighting their potential involvement in biotic stress responses. *VvNAC39* and *VvNAC44* with TC-rich repeats, STRE and W box; *VvNAC73* with W box and STRE; *VvNAC08* and *VvNAC30* with WRE3, STRE and W box; *VvNAC60* with STRE and WUN-motif; *VvNAC31* with TC-rich repeats and W box; VvNAC41 with STRE, W box and WUN-motif; finally *VvNAC61* and *VvNAC74* with STRE and WRE3.

4.5.5. Subcellular localization prediction

The subcellular localization of NAC proteins in *Vitis vinifera* L. was analyzed using the CELLO v.2.5 web database tool. Results revealed that most proteins (73%) were localized in the nucleus, while eleven proteins, namely *VvNAC08, VvNAC15, VvNAC20, VvNAC26, VvNAC30, VvNAC31, VvNAC32, VvNAC36, VvNAC37, VvNAC38,* and *VvNAC64,* were present in both the cytoplasm and nucleus.

Interestingly, *VvNAC72* was highly predicted to be located in the plastid. Moreover, *VvNAC39* and *VvNAC40* were linked to mitochondria, nuclei, and the cytoplasm, while *VvNAC05* was found in the nucleus and mitochondria. In addition, *VvNAC50* was located in the mitochondria and extracellular matrix, while *VvNAC73* was distributed in the cytoplasm and mitochondria. Furthermore, *VvNAC35* and *VvNAC51* were found in the cytoplasm, and *VvNAC52* was located in the cytoplasm and extracellular matrix.

4.5.6. Statistical Analysis

Through hierarchical clustering a statistical analysis for gene expression data was generated, by creating a tree-like structure, known as a dendrogram, the relationships between genes were visually represented.

The dendrogram formed 4 clusters of the *VvNAC* genes with similar expression profiles were grouped together in branches. Cluster 1 and cluster 2 contain low expressed genes while cluster 3 and 4 contain the high expressed genes.

The Euclidean distance were used to measure the distance between the clusters to determine their similarities in a quantitative manner based on the Euclidean formula:

$$d_{Uv} = \sqrt{(u_1 - v_1)^2 + (U_2 - v_2)^2 + (U_3 - v_3)^2}$$

The distance between the clusters 2 & 3 based on the single linkage was 4.2; the distance between the closest pair *VvNAC64&VvNAC52.*

The distance between the clusters 1 & 4 based on the single linkage was: 6; the distance between *VvNAC03&VvNAC73*.

The distance between the clusters 2 & 4 based on the single linkage was: 10.5; the distance between *VvNAC52&VvNAC73.*

While the distance between the clusters 3 & 4 based on the single linkage was: 2.27; the distance between *VvNAC30&VvNAC73.*

Therefore, the proximity of clusters 3 and 4 in terms of gene expression suggests that they may share similar regulatory mechanisms or pathways in response to *Botrytis cinerea*. Genes in these clusters may be involved in common biological processes. While the larger distance between clusters 1&4, 2&3 and 2&4 indicates dissimilarity and the genes in these clusters have distinct expression patterns. This could be due to a different regulatory mechanism or involvement in different biological processes.

4.5.7. In vitro verifications: VvNAC36 model

In *Vitis vinifera* L. a recent study has proved a remarkable regulation of the *VvNAC36* gene in *A. thaliana* in response to powdery mildew colonization. The up-regulation of *VvNAC36* 1.5-fold in the case of powdery mildew (PM) while no response was detected for salicylic acid (SA), has demonstrated the role of *VvNAC36* in response to biotic stress (Tóth et al., 2016).

<u>Transient expression</u>

Continuing Tóth's work, a simple and rapid transient transformation in tobacco leaf was performed to determine the localization of *VvNAC36* protein that might help in better understanding its role in plants.

In plants, yeast, and mammals, the GFP tag is widely used to mark proteins for localization studies. Confocal microscope was used to image GFP fusion proteins that were transiently expressed in leaves. The use of a constitutive cauliflower mosaic virus 35S promoter (CaMV 35S) enabled GFP to be seen in leaf tissues that normally have a lot of autofluorescence. The fusion proteins *VvNAC36::sGFP* and the control *35S::sGFP* were agroinfiltrated into leaves of 3 to 5-week-old *N. benthamiana* plants.

Our investigation of the subcellular localization of the *VvNAC36*::sGFP fusion protein in tobacco epidermal cells indicated that the protein was mainly present in the cytoplasmic area.

Furthermore, our analysis suggested that the *VvNAC36* protein might also be present in various organelles, as GFP was detected in different shapes. However, this observation could not be confirmed conclusively due to the limited availability of advanced biotechnology methods.

Finally, for the positive control 35S::sGFP, the detection of the green fluorescent protein showed an accumulation in both the cytoplasm and nucleus of the tobacco leaf epidermal cells.

4.6. *VvNAC36* protein

Based on CELLO Life, the localization of *VvNAC36* protein is more likely to be in the nucleus and cytoplasm with a percentage of 54.4% and 21.1% respectively.

In order to explain the *in silico* result, a transient expression has been done for *VvNAC36* gene. Thus, as a result in the case of the *VvNAC36* protein, the GFP fluorescence presented in the cytoplasmic area, while, a localization in the nucleus and the cytoplasmic area showed in the case of the control.

Consequently, comparing to the size of the nucleus of the contol, we confirm the absence of any nuclear localization for the gene *VvNAC36* and just the cytoplasmic one was approved. Despite the good results, but future studies should use more advanced techniques to deeply investigate and confirm the specific localization of this gene.

The expression analysis based on microarray data available has showed a slight up-regulation just in the early stage of infection of the ripe grape (S1), this can be explained by the presence of different CAREs responsible for biotic stress in the promoter of *VvNAC36* such as; STRE and W box.

The computational results showed a weak resistance to *B. cinerea* that occurred just in S1 and did not present in the middle or late stages of the infection. Thus, a stable transformation must be done in the future in order to check how *VvNAC36* might respond to *B. cinerea* in different stages.

5. CONCLUSIONS AND RECOMMENDATIONS

- This study aimed to highlight and interpret the functions of *VvNAC* genes in *Vitis vinifera* L., beyond their known roles in development and stress responses, using both *in vitro* and *in silico* approaches.

- Based on the grapevine genomic databases, we have characterized 74 *VvNAC* genes.

- As a result, the analyzing of the protein's lengths and number of motifs is just one aspect of a gene's potential to respond to biotic stress and it should not be the sole determinant of a gene's function.

- For example: the longest gene *VvNAC54* with a higher number of motifs (13) showed down-regulation in response to *B. cinerea*, while one of the shortest genes, the *VvNAC73* with 5 motifs showed high expression to the disease.

- While it's logical to assume that more length and motifs might make a gene more effective in responding to stress, several factors can influence the outcomes, as we have observed in the case of *VvNAC54* and *VvNAC73.*

- Continuing, with the Ka/Ks value, all genes displayed Ka/Ks values below 1. This suggests that the genes in this tandem repeat sequence are likely under strong purifying selection, and they are likely performing important functions without substantial genetic modifications.

- Moreover, in the subcellular localization analysis, it is noteworthy to mention *VvNAC72.* A gene with a specialty to be in the plastid, makes us curious and eager to investigate further.

- In addition, protein-protein interactions analysis is important because understanding these relationships, such in the case of *VvNAC08, VvNAC39* and *VvNAC44* can be valuable in biotechnological applications, such as crop improvement. It may lead to the development of strategies to enhance stress resistance or other desirable traits in grapevines based on the knowledge from *Arabidopsis thaliana.*

- The CAREs in the up-regulated genes promoters might not be unique just to *Vitis vinifera NAC genes* but they represent conserved motifs with potential involvement in biotic stress response.

- Based on the expression profiles analysis, we predicted 10 *VvNAC* genes for *B. cinerea* tolerance. Followed by a statistical analysis through hierarchical clustering for the gene's expression.

- Thus, additional experimental validation, such as quantitative real-time PCR (qRT-PCR) analysis, is required to verify the expression pattern of the predicted *VvNACs* in response to *B. cinerea.*
- In the case of the *in vitro* experiment related to *VvNAC36*, the results of the transient expression encountered with the computational tools' results.
- The stable transformation must be continued, and *in vitro* experiments should be performed to the predicted *VvNAC* genes and not just for *VvNAC36.*
- Finally, this study has increased our knowledge of grape *NAC* genes and provided insight into their functions in case of the up-regulated genes.
- Furthermore, our findings have built a robust framework for researchers to select candidates to engineer grape cultivars for enhanced tolerance against biotic stress.

6. NEW SCIENTIFIC RESULTS

1. An *in silico* approach was applied for detailed functional characterization of *VvNAC* transcription factor genes in response to *B. cinerea.*
2. We predicted 10 *VvNAC* genes (*VvNAC08, VvNAC30, VvNAC31, VvNAC36, VvNAC39, VvNAC41, VvNAC44, VvNAC61, VvNAC73,* and *VvNAC74*) for *B. cinerea* tolerance, that could be useful as reference for researchers in future grape breeding programs.
3. We have determined through a transient expression the subcellular localization of the *VvNAC36*, in the cytoplasmic area.
4. The feasibility study of *Nicotiana bentamiana* L. in transient transformation with the *VvNAC36* and the correlation between *in silico* and *in vitro* approaches were portrayed.

7. REFERENCES

1. Bailey, T. L., Boden, M., Buske, F. A., Frith, M., Grant, C. E., Clementi, L., Ren, J., Li, W. W., & Noble, W. S. (2009). MEME SUITE: tools for motif discovery and searching. Nucleic acids research, 37(Web Server issue), W202–W208. https://doi.org/10.1093/nar/gkp335.
2. Blanco-Ulate, B., Amrine, K. C. H., Collins, T. S., Rivero, R. M., Vicente, A. R., Morales-Cruz, A., Doyle, C. L., Ye, Z., Allen, G., Heymann, H., Ebeler, S. E., & Cantu, D., 2015. Transcription profiling by high throughput sequencing of berries at distinct stages of noble rot caused by Botrytis cinerea. Expression Atlas, microarray data sets, E-GEOD-67932. https://www.ebi.ac.uk/gxa/experiments/E-GEOD-67932/Results.
3. Breeze, E., Harrison, E., McHattie, S., Hughes, L., Hickman, R., Hill, C., Kiddle, S., Kim, Y. S., Penfold, C. A., Jenkins, D., Zhang, C., Morris, K., Jenner, C., Jackson, S., Thomas, B., Tabrett, A., Legaie, R., Moore, J. D., Wild, D. L., Ott, S., … Buchanan-Wollaston, V. (2011). High-resolution temporal profiling of transcripts during Arabidopsis leaf senescence reveals a distinct chronology of processes and regulation. *The Plant cell*, *23*(3), 873–894. https://doi.org/10.1105/tpc.111.083345.

4. Chen, C., Chen, H., Zhang, Y., Thomas, H. R., Frank, M. H., He, Y., & Xia, R. (2020). TBtools: An Integrative Toolkit Developed for Interactive Analyses of Big Biological Data. Molecular plant, 13(8), 1194–1202. https://doi.org/10.1016/j.molp.2020.06.009.
5. Duval, M., Hsieh, T. F., Kim, S. Y., & Thomas, T. L. (2002). Molecular characterization of AtNAM: a member of the Arabidopsis NAC domain superfamily. *Plant molecular biology*, *50*(2), 237–248. https://doi.org/10.1023/a:1016028530943.
6. Elad Y., Williamson B., Tudzynski P., Delen N. (2004). "Botrytis spp. and diseases, they cause in agricultural systems-an introduction," in *Botrytis: biology, pathology and control*. Eds. Elad Y., Williamson B., Tudzynski P., Delen N. (Dordrecht, the Netherlands: Kluwer;), 1–27. https://link.springer.com/book/10.1007%2F978-1-4020-2626-3.
7. Finn, R. D., Coggill, P., Eberhardt, R. Y., Eddy, S. R., Mistry, J., Mitchell, A. L., Potter, S. C., Punta, M., Qureshi, M., Sangrador-Vegas, A., Salazar, G. A., Tate, J., & Bateman, A. (2016). The Pfam protein families database: towards a more sustainable future. Nucleic acids research, 44(D1), D279–D285. https://doi.org/10.1093/nar/gkv1344.
8. Goodstein, D. M., Shu, S., Howson, R., Neupane, R., Hayes, R. D., Fazo, J., Mitros, T., Dirks, W., Hellsten, U., Putnam, N., & Rokhsar, D. S. (2012). Phytozome: a comparative platform for green plant genomics. Nucleic acids research, 40(Database issue), D1178–D1186. https://doi.org/10.1093/nar/gkr944.
9. Gusain, P, Uniyal, DP, Joga, R. (2021). Conservation and sustainableuse of medicinal plants. In: Egbuna, C, Mishra, AP, Goyal, MR,editors. Preparation of phytopharmaceuticals for the manage-ment of disorders. London: Academic Press; 2021. p. 409–27. https://doi:10.1016/B978-0-12-820284-5.00026-5.
10. Haile, Z. M., Pilati, S., Sonego, P., Malacarne, G., Vrhovsek, U., Engelen, K., Tudzynski, P., Zottini, M., Baraldi, E., & Moser, C. (2017). Molecular analysis of the early interaction between the grapevine flower and Botrytis cinerea reveals that prompt activation of specific host pathways leads to fungus quiescence. *Plant, cell & environment*, *40*(8), 1409–1428. https://doi.org/10.1111/pce.12937.
11. Hoshikawa, K., Fujita, S., Renhu, N., Ezura, K., Yamamoto, T., Nonaka, S., Ezura, H., & Miura, K. (2019). Efficient transient protein expression in tomato cultivars and wild species using agroinfiltration-mediated high expression system. *Plant cell reports*, *38*(1), 75–84. https://doi.org/10.1007/s00299-018-2350-1.
12. Ibraheem, O., Botha, C. E., & Bradley, G. (2010). *In-silico* analysis of cis-acting regulatory elements in 5' regulatory regions of sucrose transporter gene families in rice (Oryza sativa Japonica) and Arabidopsis thaliana. *Computational biology and chemistry*, *34*(5-6), 268–283. https://doi.org/10.1016/j.compbiolchem.2010.09.003.
13. Ison, J., Kalas, M., Jonassen, I., Bolser, D., Uludag, M., McWilliam, H., Malone, J., Lopez, R., Pettifer, S., & Rice, P. (2013). EDAM: an ontology of bioinformatics operations, types of data and identifiers, topics and formats. Bioinformatics (Oxford, England), 29(10), 1325–1332. https://doi.org/10.1093/bioinformatics/btt113.
14. Ju, Y. L., Yue, X. F., Min, Z., Wang, X. H., Fang, Y. L., & Zhang, J. X. (2020). *VvNAC17*, a novel stress-responsive grapevine (*Vitis vinifera L.*) NAC transcription

factor, increases sensitivity to abscisic acid and enhances salinity, freezing, and drought tolerance in transgenic Arabidopsis. *Plant physiology and biochemistry: PPB*, *146*, 98–111. https://doi.org/10.1016/j.plaphy.2019.11.002.

15. Kaur, G., & Pati, P. K. (2016). Analysis of cis-acting regulatory elements of Respiratory burst oxidase homolog (Rboh) gene families in Arabidopsis and rice provides clues for their diverse functions. *Computational biology and chemistry*, *62*, 104–118. https://doi.org/10.1016/j.compbiolchem.2016.04.002.
16. Keller, M., Viret, O., & Cole, F. M. (2003). Botrytis cinerea Infection in Grape Flowers: Defense Reaction, Latency, and Disease Expression. *Phytopathology*, *93*(3), 316–322. https://doi.org/10.1094/PHYTO.2003.93.3.316.
17. Kim, Y. S., Kim, S. G., Park, J. E., Park, H. Y., Lim, M. H., Chua, N. H., & Park, C. M. (2006). A membrane-bound NAC transcription factor regulates cell division in Arabidopsis. *The Plant cell*, *18*(11), 3132–3144. https://doi.org/10.1105/tpc.106.043018.
18. Ko, J. H., Yang, S. H., Park, A. H., Lerouxel, O., & Han, K. H. (2007). *ANAC012*, a member of the plant-specific NAC transcription factor family, negatively regulates xylary fiber development in Arabidopsis thaliana. *The Plant journal: for cell and molecular biology*, *50*(6), 1035–1048. https://doi.org/10.1111/j.1365-313X.2007.03109.x.
19. Krzywinski, M., Schein, J., Birol, I., Connors, J., Gascoyne, R., Horsman, D., Jones, S. J., & Marra, M. A. (2009). Circos: an information aesthetic for comparative genomics. Genome research, 19(9), 1639–1645. https://doi.org/10.1101/gr.092759.109.
20. Le Hénanff, G., Profizi, C., Courteaux, B., Rabenoelina, F., Gérard, C., Clément, C., Baillieul, F., Cordelier, S., & Dhondt-Cordelier, S. (2013). Grapevine *NAC1* transcription factor as a convergent node in developmental processes, abiotic stresses, and necrotrophic/biotrophic pathogen tolerance. *Journal of experimental botany*, *64*(16), 4877–4893. https://doi.org/10.1093/jxb/ert277.
21. Lescot, M., Déhais, P., Thijs, G., Marchal, K., Moreau, Y., Van de Peer, Y., Rouzé, P., & Rombauts, S. (2002). PlantCARE, a database of plant cis-acting regulatory elements and a portal to tools for in silico analysis of promoter sequences. *Nucleic acids research*, *30*(1), 325–327. https://doi.org/10.1093/nar/30.1.325.
22. Li, L., He, Y., Zhang, Z., Shi, Y., Zhang, X., Xu, X., Wu, J. L., & Tang, S. (2021). OsNAC109 regulates senescence, growth and development by altering the expression of senescence- and phytohormone-associated genes in rice. *Plant molecular biology*, *105*(6), 637–654. https://doi.org/10.1007/s11103-021-01118-y.
23. Lindbo J. A. (2007). High-efficiency protein expression in plants from agroinfection-compatible Tobacco mosaic virus expression vectors. *BMC biotechnology*, *7*, 52. https://doi.org/10.1186/1472-6750-7-52.
24. Masri, R., & Kiss, E. (2023). The role of NAC genes in response to biotic stresses in plants. *Physiological and Molecular Plant Pathology*, 102034. https://doi.org/10.1016/j.pmpp.2023.102034.
25. Pezet, R., Viret, O., Perret, C. and Tabacchi, R. (2003), Latency of *Botrytis cinerea* Pers.: Fr. and Biochemical Studies During Growth and Ripening of Two Grape Berry Cultivars,

Respectively Susceptible and Resistant to Grey Mould. Journal of Phytopathology, 151: 208-214. https://doi.org/10.1046/j.1439-0434.2003.00707.x.

26. Priest, H. D., Filichkin, S. A., & Mockler, T. C. (2009). Cis-acting regulatory elementss in plant cell signaling. *Current opinion in plant biology*, *12*(5), 643–649. https://doi.org/10.1016/j.pbi.2009.07.016.
27. Prusky, D., Alkan, N., Mengiste, T., & Fluhr, R. (2013). Quiescent and necrotrophic lifestyle choice during postharvest disease development. *Annual review of phytopathology*, *51*, 155–176. https://doi.org/10.1146/annurev-phyto-082712-102349.
28. Puranik, S., Sahu, P. P., Srivastava, P. S., & Prasad, M. (2012). NAC proteins: regulation and role in stress tolerance. *Trends in plant science*, *17*(6), 369–381. https://doi.org/10.1016/j.tplants.2012.02.004.
29. Rhee, S. Y., & Mutwil, M. (2014). Towards revealing the functions of all genes in plants. *Trends in plant science*, *19*(4), 212–221. https://doi.org/10.1016/j.tplants.2013.10.006.
30. Shao, H., Wang, H., & Tang, X. (2015). NAC transcription factors in plant multiple abiotic stress responses: progress and prospects. *Frontiers in plant science*, *6*, 902. https://doi.org/10.3389/fpls.2015.00902.
31. Sperotto, R. A., Ricachenevsky, F. K., Duarte, G. L., Boff, T., Lopes, K. L., Sperb, E. R., Grusak, M. A., & Fett, J. P. (2009). Identification of up-regulated genes in flag leaves during rice grain filling and characterization of *OsNAC5*, a new ABA-dependent transcription factor. *Planta*, *230*(5), 985–1002. https://doi.org/10.1007/s00425-009-1000-9.
32. Szklarczyk, D., Kirsch, R., Koutrouli, M., Nastou, K., Mehryary, F., Hachilif, R., Gable, A. L., Fang, T., Doncheva, N. T., Pyysalo, S., Bork, P., Jensen, L. J., & von Mering, C. (2023). The STRING database in 2023: protein-protein association networks and functional enrichment analyses for any sequenced genome of interest. Nucleic acids research, 51(D1), D638–D646. https://doi.org/10.1093/nar/gkac1000.
33. Toth, Z., Winterhagen, P., Kalapos, B., Su, Y., Kovacs, L., & Kiss, E. (2016). Expression of a Grapevine *NAC* Transcription Factor Gene Is Induced in Response to Powdery Mildew Colonization in Salicylic Acid-Independent Manner. *Scientific reports*, *6*, 30825. https://doi.org/10.1038/srep30825.
34. Veloso, J., & van Kan, J. (2018). Many Shades of Grey in Botrytis-Host Plant Interactions. *Trends in plant science*, *23*(7), 613–622. https://doi.org/10.1016/j.tplants.2018.03.016.
35. Wang, N., Zheng, Y., Xin, H., Fang, L., & Li, S. (2013). Comprehensive analysis of NAC domain transcription factor gene family in Vitis vinifera. *Plant cell reports*, *32*(1), 61–75. https://doi.org/10.1007/s00299-012-1340-y.
36. Yu, C. S., Lin, C. J., & Hwang, J. K. (2004). Predicting subcellular localization of proteins for Gram-negative bacteria by support vector machines based on n-peptide compositions. Protein science: a publication of the Protein Society, 13(5), 1402–1406. https://doi.org/10.1110/ps.03479604.

8. PUBLICATIONS

- Masri, R., & Kiss, E. (2023). The role of *NAC* genes in response to biotic stresses in plants. *Physiological and Molecular Plant Pathology*, 102034. https://doi.org/10.1016/j.pmpp.2023.102034 IF:2.7

- Masri, R., & Kiss, E. (2023). An in silico approach for characterizing *Vitis vinifera NAC* promoters in response to *Botrytis cinerea.* (Accepted for publication in the journal Current Plant Biology).