Article

Identification and Characterization of *TALE* Homeobox Genes in the Endangered Fern *Vandenboschia speciosa*

Mercedes Ruiz-Estévez, Mohammed Bakkali [ID]**, Rubén Martín-Blázquez and Manuel A. Garrido-Ramos ***

Departamento de Genética, Facultad de Ciencias, Universidad de Granada, 18071 Granada, Spain;
mercy_canaria@hotmail.com (M.R.-E.); mbakkali@ugr.es (M.B.); rmblazquez@ugr.es (R.M.-B.)
* Correspondence: mgarrido@ugr.es; Tel.: +34-958-249-710

Abstract: We report and discuss the results of a quantitative reverse transcription polymerase chain reaction (qRT-PCR) analysis of the expression patterns of seven three amino acid loop extension (*TALE*) homeobox genes (four *KNOTTED-like homeobox* (*KNOX*) and three *BEL1-like homeobox* (*BELL*) genes) identified after next generation sequencing (NGS) and assembly of the sporophyte and gametophyte transcriptomes of the endangered fern species *Vandenboschia speciosa*. Among the four *KNOX* genes, two belonged to the *KNOX1* class and the other two belonged to the *KNOX2* class. Analysis of the deduced amino acid sequences supported the typical domain structure of both types of TALE proteins, and the homology to TALE proteins of mosses, lycophytes, and seed plant species. The expression analyses demonstrate that these homeodomain proteins appear to have a key role in the establishment and development of the gametophyte and sporophyte phases of *V. speciosa* lifecycle, as well as in the control of the transition between both phases. *Vandenboschia speciosa* VsKNAT3 (a KNOX2 class protein) as well as VsBELL4 and VsBELL10 proteins have higher expression levels during the sporophyte program. On the contrary, one *V. speciosa KNOX1* protein (VsKNAT6) and one KNOX2 protein (VsKNAT4) seem important during the development of the gametophyte phase. *TALE* homeobox genes might be among the key regulators in the gametophyte-to-sporophyte developmental transition in regular populations that show alternation of generations, since some of the genes analyzed here (*VsKNAT3*, *VsKNAT6*, *VsBELL4*, and *VsBELL6*) are upregulated in a non-alternating population in which only independent gametophytes are found (they grow by vegetative reproduction outside of the range of sporophyte distribution). Thus, these four genes might trigger the vegetative propagation of the gametophyte and the repression of the sexual development in populations composed of independent gametophytes. This study represents a comprehensive identification and characterization of *TALE* homeobox genes in *V. speciosa*, and gives novel insights about the role of these genes in fern development.

Keywords: *TALE* homeobox genes; *KNOX* genes; *BELL* genes; fern; *Vandenboschia speciosa*; expression patterns; qRT-PCR; transcriptome

1. Introduction

The phylogenetic position of ferns as the sister lineage of seed plants makes the study of developmental genes in these vascular plants highly useful, and necessary to open new insights on plant evolutionary developmental biology [1–7]. *Vandenboschia speciosa* is a rare vulnerable European-Macaronesian endemism, the only representative of a genus which has a primarily tropical distribution, restricted to disjoined regions of the European Atlantic coast and the Macaronesian islands (Canaries, Madeira and Azores). These regions are tertiary flora refuges harbouring relic populations

that survived the glacial cycles. The two phases of *V. speciosa* life cycle are perennial, and can reproduce by vegetative propagation [8]. The "floaty" sporophyte (fronds made of a translucent single layer of cells) is rhizomatous, and can spread by fragmentation of the rhizome. The gametophyte, unlike a heart-shaped prothallus, is epigeous and narrowly filamentous with specialized asexual propagules (gemmae). While the sporophyte is adapted to growth in areas with low incidence of light and constant moisture, the gametophyte can live in a wider range of habitats, including less humid and darker ones, and can survive in very sheltered stable microsites, such as caves and deep crevices [8]. Such habitats can provide a microclimate and a stable environment for survival during long periods of independent gametophytes, outside the range of sporophyte distribution [8]. The prolonged gametophyte survival could be facilitated by a low metabolic rate and an ability to make efficient use of the limited available light, thanks to morphological and physiological adaptations [9,10]. In the warmer climatic conditions in the South of the Iberian Peninsula, the Azores, the Madeira, and the Canaries, this species usually undergoes a normal fern lifecycle of two free-living generations, but, as one goes further north and east of Europe, the sporophyte generation becomes increasingly rare [8]. However, one out of seven populations of this species located in the south of the Iberian Peninsula shows only the gametophyte phase. All these features make this species an attractive species for the analysis of different developmental and evolutionary topics, including the ones with important implications for conservation biology. In this context, it is highly interesting to unravel and understand the genetics behind the control of the alternation of the two fern lifecycle phases by characterizing and analyzing homeobox genes, an issue for which the analysis of gametophyte independent populations of *V. speciosa* might represent an important advance.

Homeobox genes encode a typical DNA-binding domain known as homeodomain (HD) that characterizes a large family of transcription factors. In plants, fourteen distinct classes of homeobox genes are found [11–13]. The three amino acid loop extension (*TALE*) homeobox gene superclass is characterized by encoding three extra residues (proline–tyrosine–proline) within the loop between the helix 1 and helix 2 of the homeobox-encoded homeodomain [11–17]. The *TALE* homeobox genes are very ancient, being present in a wide range of eukaryotic kingdoms with five classes in animals and two classes in plants, the *KNOTTED-like homeobox* (*KNOX*) and *BEL1-like homeobox* (*BELL*) genes. Both plant *TALE* genes encode large bipartite domains upstream of the homeodomain, termed *KNOX* and *BEL* domains, respectively. In addition, the *KNOX* genes encode the *ELK* domain between the KNOX and the HD domains. The *BELL* genes additionally encode also the SKY domain and the ZIBEL domains. BELL and KNOX proteins are structurally and functionally related, and have been shown to interact through their BELL and KNOX domains, respectively [14,16,18–23]. BELL and KNOX transcription factors appear to be key for the distinctive gametophytic and sporophytic developmental programs, as well as for the gametophyte-to-sporophyte developmental transition that characterizes plants [21–25]. The diversification of KNOX/BELL modules during land plant evolution facilitated the evolution of ever more complex diploid sporophyte body plans [21,23,25].

Divergence between *KNOX* and *BELL* gene lineages occurred before the split between red and green algae [22,25]. Green algae have a single *KNOX* and a single *BELL* gene. Prior to the origin of land plants, a gene duplication in an ancestral *KNOX* gene generated two classes of *KNOX* genes, class I (*KNOX1*) and class II (*KNOX2*) [11,22,25–27], increasing fundamental structural and developmental challenges related to the transition to land [22,28–31]. In fact, the *KNOX* genes have experienced two major events of expansion during the evolution of plants, one from algae to moss and the other during the transition from lycophytes to angiosperms [22]. There are five *KNOX* genes in bryophytes (three *KNOX1* and two *KNOX2*) and four *KNOX* genes in lycophytes (two *KNOX1* and two *KNOX2*), while there has been a diversification of *KNOX* genes in vascular plants [11,14,18,22,25,32]. In *Arabidopsis thaliana*, there are eight *KNOX* genes (four *KNOX1* and four *KNOX2* genes) [22]: *SHOOT MERISTEMLESS* (*STM*), *KNAT1*, *KNAT2*, and *KNAT6* are *KNOX1* genes, while *KNAT3*, *KNAT4*, *KNAT5*, and *KNAT7* are *KNOX2* genes. There are nineteen *KNOX* genes in poplar and fourteen genes in rice [11,23,25]. The diversification of *BELL* genes has also accompanied the evolution of land plants.

There are four *BELL* genes in mosses and thirteen *BELL* genes in *A. thaliana* and in potato, while there are fifteen in poplar and twelve in rice [11,17,23,25,33].

In general, there is unfortunately little information about developmental genes in ferns. Furthermore, the genomic data from fern species are not extensive [34], and although new insights are appearing from recent efforts on the genomics of ferns [35–37], only a few studies focused on the analysis of individual *KNOX* genes from fern species [27,38–40]. In this paper, we report and discuss the results of a quantitative PCR analysis of the expression patterns of four *KNOX* and three *BELL* genes identified after next generation sequencing (NGS) and assembly of the transcriptomes of the sporophyte and of the gametophyte of the endangered fern species, *V. speciosa*. The results are compared between genes, phases, and populations, in order to investigate the different expression patterns between populations with and without alternation of generations, and infer on the role of those genes in fern development.

2. Materials and Methods

2.1. Sample Collection

We collected sporophyte, spore filled sporangia, and gametophyte of *V. speciosa* from the population of Valdeinfierno (VALD), located at the Natural Park of Alcornocales (Cádiz, Spain). For comparisons, sporophyte and gametophyte of *V. speciosa* from a second population (Canuto de Ojén Quesada (OJEN)) were also collected. VALD and OJEN are two populations with a regular fern lifecycle of two free-living generations, gametophyte and sporophyte. In addition, gametophytes of this species from La Almoraima (ALMO) were also collected. The ALMO population is composed only of independent gametophytes. The sporophytes were frozen in liquid nitrogen immediately at the moment they were collected. The patches with gametophytes were taken in a Petri dish with soil to the laboratory, where, under a binocular microscope, the filaments were separated from the soil and from other plant species, cleaned, and frozen in liquid nitrogen. All samples were stored at $-80\ ^{\circ}$C for RNA extractions.

2.2. Next Generation Sequencing, Transcriptome Assembly, Annotation, and Comparative In Silico Analysis of Gene Expression Levels

For transcriptome sequencing, RNA was isolated from five gametophytes and five sporophytes from VALD (April) using Spectrum Plant Total RNA Kit (Sigma, Madrid, Spain). Pools of sporophyte and gametophyte RNA were generated from the two sets of five individual RNAs and separated. Next generation sequencing of the two samples was carried out (Macrogen Inc., Seoul, Korea) based on the Illumina HiSeq 2000 Paired-end approach (Illumina Inc., Seoul, South Korea).

After assessment of the quality of the resulting sequencing reads, using FastQC [41], ABySS [42] was used to carry out a first multi k-mer transcriptome assembly step using all the sequencing reads in FASTQ format. The k-mers included all the odd numbers from 19 to 95.

A first set of one transcriptome per k-mer was thus generated, and a second assembly step was then carried out in order to merge all these transcriptomes into one, filter the sequences into the merged file to remove redundancy, and extend the remaining non-redundant sequences. This second step was based on the Trans-ABySS pipeline [43], and allowed us to further assemble the reference transcriptome and get the transcripts as lengthy as possible. A final assembly step was also added in order to further extend the assembled transcripts, this time using the CAP3 software [44].

The sequences of the resulting FASTA file were then BLASTed against a locally built database with all the *Arabidopsis*, fern, and moss proteins from the NCBI database. For that, local BLASTX searches were carried out [45], and the hits were considered positive at a 10^{-6} *e*-value threshold. Blast2GO software [46] was then used for functional annotation of the Basic Local Alignment Search Tool (BLAST) positive sequences.

For comparison of the expression levels, both BLAST positive and BLAST negative sequences were taken as reference transcriptome in annotated FASTA format. The raw sequencing reads (the FASTQ files) from the libraries of the sporophytes and the gametophytes were then separately aligned to the reference sequences using BWA [47]. The number of reads that aligned to each reference sequence was counted using the *htseq-count* script of the HTSeq program [48]. For fold change-based comparison of the number of reads aligned to each sequence between the two libraries, and we carried out normalization of the read counts into Reads per Kilobase of the reference sequence per Million reads in the library (RPKM). Assessment of the statistical significance of the differences was based on a contingency χ^2-test strategy and false discovery rate correction [49].

2.3. Sequence Analysis

BioEdit software (version 7.1.3) [50] was used to deduce the amino acid sequences from the NGS assembled contigs that were earlier annotated as sequences corresponding to homeobox genes. BLAST [45] was then used to search in the protein databases with the deduced amino acid sequences as queries. Sequences from other plants were collected from gene databases. Multiple alignments of deduced amino acid sequences were performed using ClustalX (default parameters), which implement the ClustalW algorithm [51], and adjusted manually afterwards.

2.4. Expression Analysis

2.4.1. RNA Extraction and Complementary DNA Synthesis

Total RNA was extracted from sporophytes and gametophytes using the Spectrum Plant Total RNA Kit (Sigma). After RNA extraction, the samples were submitted to a DNase I treatment (REAL Star Kit, Durviz, Paterna, Spain), to ensure the absence of contaminating genomic DNA traces. RNA quantity and purity were measured using Tecan'sInfinite 200 NanoQuant. For each sample, 100 ng RNA was retrotranscribed into complementary DNA (cDNA) using the combination of random and oligo deoxythymidine (dT) primers of the PrimeScript RT reagent—Perfect Real Time-Kit (Takara) and following manufacturer's recommendations. For each analysis, a negative control devoid of reverse transcriptase was included to test for contaminating DNA.

2.4.2. Target and Housekeeping Genes

The homeobox and housekeeping (*HKG*) genes of *V. speciosa* were identified using BLASTN searches of other plant homeobox and *HKG* sequences against our local *V. speciosa* transcriptome library. The *HKG* sequences searched were those provided by Le Bail et al. [52]. Primers were designed using Primer3 software [53] (Table S1; Supplementary material). We tested their specificity and reliability by PCR amplification and posterior sequencing. PCR reactions mixture contained 30 ng complementary DNA (cDNA), 200 mM deoxynucleotide triphosphates (dNTPs), 10 mM each forward and reverse primers, 2 mM $MgCl_2$, and 1 unit of Horse-Power-Taq DNA polymerase (Canvax, Granada, Spain) in a final volume of 20 μL. PCR assays were run in a GeneAmp PCR system 2700 thermocycler (Applied Biosystems, Madrid, Spain), and PCR conditions were as follows: an initial denaturation at 94 °C for 5 min; and 35 cycles of 94 °C for 30 s, 55 °C for 20 s, 72 °C for 20 s; and final extension of 72 °C for 4 min. PCR products were visualized after electrophoresis in 1.5% agarose gel, and cleaned with illustra GFXPCR DNA and Gel Band Purification Kit (GE Healthcare, Madrid, Spain). The PCR products were sequenced at the Centro de Instrumentación Científica (University of Granada), and analyzed using BioEdit software (version 7.1.3) [50].

2.4.3. Relative Gene Expression Quantification

After ensuring the specificity of the target and *HKG* primers, we determined the most stable *HKG*s in our samples through carrying out a geNorm analysis [54]. Standard curve analysis was used to determine the efficiency of the selected *HKG*s. We then estimated the relative expression

level of the homeobox genes in the different samples by means of quantitative reverse transcription polymerase chain reaction (qRT-PCR). The reaction mixtures contained 5 μL 2X SensiMix SYBR Mastermix (SensiMix SYBR Kit, Bioline, Barcelona, Spain), 0.35 μM each forward and reverse primer, and 5 ng cDNA, in a final volume of 15 μL. We amplified the same calibrator sample (comprising cDNA synthesized from RNA of different samples) in each run to ensure that the data resulting from the experimental samples were comparable. qRT-PCR assays were run in a Chromo4 Real Time PCR thermocycler (BioRad, Alcobendas, Madrid, Spain), and PCR conditions were as follows: an initial denaturation at 95 °C for 10 min; 40 cycles of 94 °C for 30 s, 55 °C for 20 s, 72 °C for 20 s; and a melting curve step to check the specificity of the reaction. We included a negative control without cDNA to ensure that the reagents were free of contaminating DNA. There was no amplification from this sample, and therefore, the test allowed us to determine that the primers (or any PCR reagent) were not contaminated, nor had the primers formed dimers with each other. Where possible, reagents were combined in master mix solutions to minimize the number of manipulations, and each sample was amplified in triplicate. Opticon Monitor v3.1. software (BioRad) was used to export the qRT-PCR raw data from the Chromo4 instrument and relative quantification (RQ) values of the transcripts were obtained following the "Efficiency calibrated mathematical method for the relative expression ratio in real-time PCR" (Roche Applied Science, Technical Note No. LC 13/2001). This method allows you to use more than one *HKG*s to normalize the C_t values and calculate the RQs, employing the geometric average of the Q the *HKG*s used. We identified the specificity of the amplified products by sequencing the qRT-PCR products.

2.4.4. Statistical Analyses

All homeobox gene RQs failed to fit a normal distribution (tested by the Shapiro–Wilk's test), and thus, the nonparametric Mann–Whitney *U* test was used to analyze them. All these analyses were performed using the STATISTICA 8.0 software (TIBCO Software Inc., Palo Alto, CA, USA) [55].

3. Results

3.1. Identification and Sequence Analysis of TALE Homeobox Genes Contigs in V. Speciosa

Examination of the two NGS assembled transcriptomes of *V. speciosa* revealed 11 contigs corresponding to 11 putative *KNOX* genes: 3 belonged to the *KNOX1* class and 8 belonged to the *KNOX2* class. In addition, 11 putative BEL1-like type genes were detected. We selected four *KNOX* genes (two of each class) and three *BELL* genes for further analysis. In the case of *KNOX* genes, the selected genes were identified, according their origin and homology, as Vs*KNOX1* and Vs*KNOX2* class gene and named *VsKNAT1*, *VsKNAT3*, *VsKNAT4*, and *VsKNAT6*, according to their homology to different types of *KNOX* (*KNAT*) genes in *A. thaliana*. In the case of *BELL* genes, they were named as *VsBELL4*, *VsBELL6*, and *VsBELL10*, according to their homologous partners in *A. thaliana*. The sequences (Figure S1; Supplementary material) were deposited in the publicly accessible repository of the European Nucleotide Archive (ENA) of the European Bioinformatics Institute (EMBL-EBI).

The identity of each VsKNAT and VsBELL amino acid deduced sequences to its homologous proteins from different plant groups, as revealed by BLAST search, is indicated in Table 1. Homologous KNOX sequences found in GeneBank/EMBL databases [56,57] belonged to the moss *Physcomitrella patens*, the fern species *Ceratopteris richardii* and *Elaphoglossum peltatum*, and to several spermatophyte species (Figure 1 and Table 1). Homologous BELL sequences found in GeneBank/EMBL databases [56,57] belonged to the lycophyte species *Selaginella moellendorffii* and *Selaginella kraussiana*, and to several spermatophyte species (Figure 2 and Table 1).

Sequence alignments show a typical KNOX structure for the deduced VsKNAT1 and VSKNAT4 amino acid sequences of *V. speciosa*, which consist of the large bipartite KNOX domain, the ELK domain, and the homeodomain, the latter characterized by having three extra residues, PYP, in the loop between helix 1 and helix 2 of the homeodomain (Figure 1). However, the sequences obtained

for VsKNAT3 and VsKNAT6 only covered part of the bipartate KNOX domain, and did not include the other domains. The sequence alignments (Figure 1) clearly defined the VsKNAT1 and VsKNAT6 as *KNOX1* class genes, while VsKNAT3 and *VsKNAT4* belonged to the *KNOX2* class, as previously identified by the BLAST search. Sequence alignments also indicated that the deduced VsBELL amino acid sequences have a typical structure consisting of the ZYBEL domain, the SKY domain, the BEL domain, and the three amino acid loop extension homeodomain (Figure 2).

Table 1. Sequence identity of each *Vandenboschia speciosa KNAT* (VsKNAT) and *BELL* (VsBELL) proteins to different KNOX and BELL proteins from different plant species, according to the Basic Local Alignment Search Tool BLAST analysis.

	Ferns [1]	Lycophytes [2]	Moss [3]	Spermatophytes
VsKNAT1	63–69%	54%	–	44–60%
VsKNAT3	49%	43–45%	39–40%	48–88%
VsKNAT4	43%	41–46%	39–45%	46–49%
VsKNAT6	70–75%	–	–	51–65%
VsBELL4	–	40–59%	46–48%	37–43%
VsBELL6	–	45–52%	45–46%	45–52%
VsBELL10	–	47%	48–50%	47–54%

[1] *Ceratopteris richardii* and *Elaphoglossum peltatum*; [2] *Selaginella moellendorffi* and *Selaginella kraussiana*; [3] *Physcomitrella patens*.
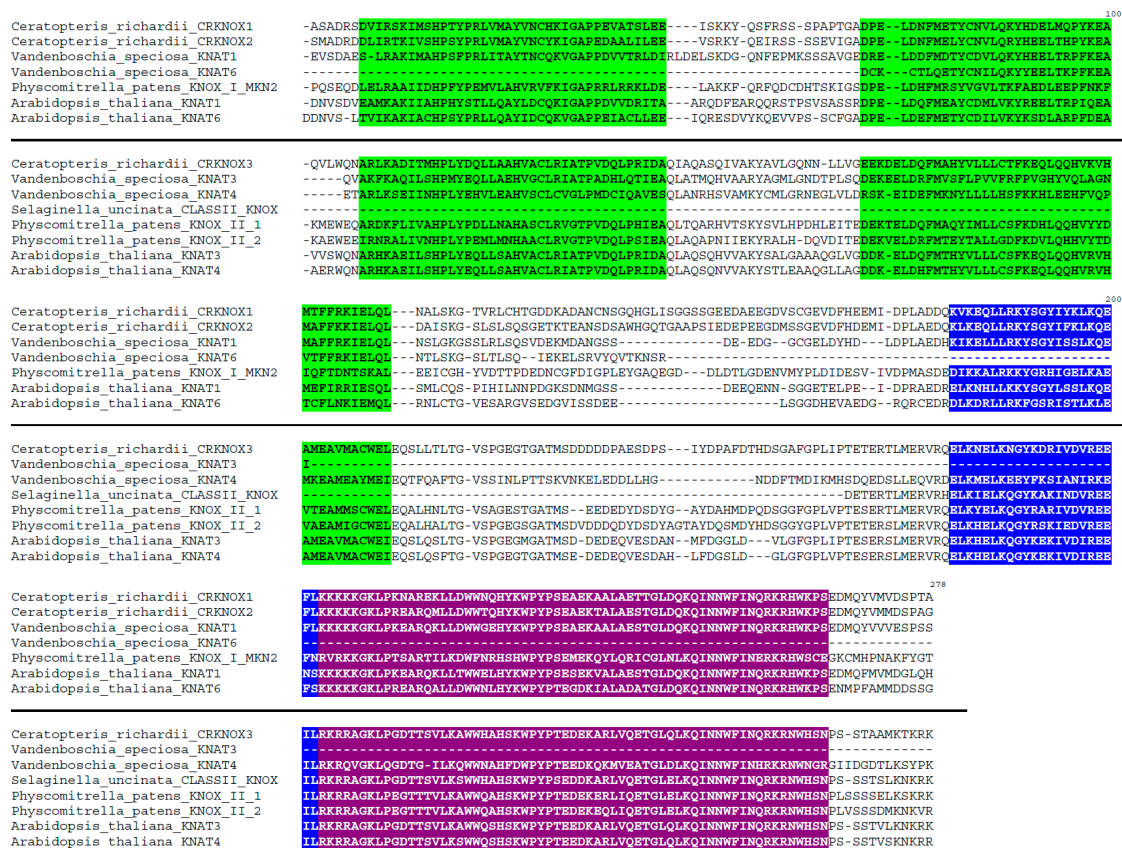
Figure 1. Alignment of the deduced amino acid sequences of the *Vandenboschia speciosa KNOTTED-like homeobox* 1 and 2 (VsKNOX1 and VsKNOX2) proteins with those of other related KNOX proteins of different plant species: the fern *Ceratopteris richardii*, the moss *Physcomitrella patens*, and the seed plant *Arabidopsis thaliana*. The sequences above the line belong to the class KNOX1 of KNOX proteins, whereas the sequences the sequences below the line belong to the class KNOX2. Green boxes mark the KNOX domains, the blue box marks the ELK domain, and the purple box marks the homeodomain.

```
                                                                                                    100
Vandenboschia_speciosa_Bel4    --------------------------RQGLSLSLSPQQHPVQMQYYSVHPNDSSVSSCAGLAHVNADQNNETRGGRFPSKWIGNLSQFRALSRDGVLATGFQNP
Vandenboschia_speciosa_Bel6    SMLDNGVALGLDSDNSHGAASSRQGLSLSLSPQQHPVQMQYYSVHPNDSSVSSCAGLAHVNADQNNETRGGRFPSKWIGNLSQFRALSRDGVLATGFQNP
Vandenboschia_speciosa_Bel10   --------------------PGLGHVAVEDSANRGERAPSKWIGSLSHFRACSRDGLLGSGFQNSTNDINSRHMHMES
Selaginella_moellendorffii     LQAGHRLEMDNSHQGAAAGPGVHQSLSLSLSN------------QNAQLHAREPGMNSPAGSQDAKRGVGKLS----GDLSHFRTSPRDGGASYVNLPS
Phoenix_dactylifera_Bel1_like  TQLGILNGQELSLQHSNVSTMQNQGLSLSLST-----QIP----VPSFQYRPATSDLS-FFGPHQSTTGNSG-------SCRDDESRNKNMHAN-----
Theobroma_cacao_Bel1_like      EMLVMHPMGGPASILHGGQSLQQGLSLSLGT-----QIPSGIQMPSMPYRNPNSGFASFLSPNQSLTGEGGSR----NSSSRDEQSRNAEYLAPGFSGG
Populus_euphratica_Bel1_like   EMLVMHSVGASSVLHGAQSLQGLSLSLGT-----QIQSGIQMPSISYRNTNPGLTSFLSPNSSITGDGGGR----NGSSRDDQSRNIEYLTPGFSGG
Vitis_vinifera_Bell_like       SSGGAADVTSQLGLRRQVVSPTQQGLSLSLSP--------------HQPTYRSVPGEHDIQVQQQPP------------VQAISPTSGDDMRVSG----
Solanum_tuberosum_Bell-like_5  LASQLAFQRPIPTPQHRQQQQQQQGLSLSLSP----QLQQQISFNNNISSSSPRTNNVT------IRGTL-------DGSSS----------
Arabidopsis_thaliana_BLH1      SLYGSHQVDPTHQQAACETPRAQQGLSLTLSS-----Q-QQQHQ-PIHVGFGSGHGE-DIRVGSGST----------------------
Hordeum_vulgare_vulgare        EAFQAYGNAASSFGGGGASVVEGQGLSLSLSP-----SLQQLEMAKQAEELRVRDGVLYFNRQQQ--------------QQQQQQAASVQQLPMALHGQ
Triticum_aestivum_Bel1_like    EAFQAYGNAASSFGGGGAGVVEGQGLSLSLSP-----SLQQLEMAKQAEELRVRDGVLYFNRQQQ--------------QQQQ--APSVQQLPMALHGQ

                                                                                                    200
Vandenboschia_speciosa_Bel4    -------LDAKFSHSGTLGLLRNSKYLRAAQQVLDEFCH-VGRSEFMSLQPNQADQISQAWMKRGFTRVVEDKIDDQAIHPESLDQNRSPVSIS
Vandenboschia_speciosa_Bel6    MNDTSSRHRQPDA--NVPTLVGFGGPLSCSKYLPAQEVLDEVVS-VQR----GIKGSS-SRHSKLPESWMASSSFLDGAFARERSLPEILSVAKEGISC
Vandenboschia_speciosa_Bel10   MNDTSSRHRQPDANVNVSTVVGPGGLLTGSKYLKAAQQLLDEVVS-VAR----VVKSSS-SKHLKL-QAWMAPNSFMDGSFARNGSLPEAVATAKEGTAG
Selaginella_moellendorffii     SAGTKHSYFDVAG-PGPSAVSNSFSFVSGSRYLRAAQQLLDEVCS-VGR----GLKQSSKSKGSQQGLGGQSSPAAVTSSLHKEAALTEN-SVKSEITIG
Phoenix_dactylifera_Bel1_like  --------LSAYG------VSSLTNPISNSKYLKAAQELLDEVVN-VRK----ALKDKADKSQSVHTSAGTAGCKDTDGATKSEG------------MP
Theobroma_cacao_Bel1_like      NQDSNKGDMSAYG------MSSIARTIPNSKYLKAAQQLLDEVVN-VRK----ALKQPDGEKNQSSQDR-MKSSTDDDRGSKN-------------VP
Populus_euphratica_Bel1_like   NHDVNKGDLSPYG------MNNIARTIPNSKYLKAAQQLLDEVVN-VRK----AIKQPDKEKNQTTSEHGLNSAKDGDSGSKN-------------GA
Vitis_vinifera_Bell_like       -GSSSTASAVSNG------ISGMQSVLLGSKYLKAAQQLLDEVAN-VGK----GIKTDSGEETKEREKVNTISVAASTGEALS----------
Solanum_tuberosum_Bell-like_5  ---------------------NMV-LGSKYLKAAQELLDEVVNIVGKSIKGDDQKKDN-SMNKESMPLASDVNTNSSGGG-----------------
Arabidopsis_thaliana_BLH1      -GSGV--------------TNGIANLVSSKYLKAAQQLLDEVVN-ADS----DDMNAKSQLFSSKKGSCGNDKPVGESSAGA-----------GGEGS
Hordeum_vulgare_vulgare        VGSMGGQQLHVGYG------PAGVAGVLRNSKYTRAAQELLDEFCS-VGRGQ--TIKGGGR--GGSSSNPNASKGGPSSSGAAQ------------------
Triticum_aestivum_Bel1_like    VGAMGQQLHVGYG------PAGVAGVLRNSKYTRAAQELLDEFCS-VGRGQ--TIKGGGRGAGGSSSNPNASKGGPSSSGAGQ------------------

                                                                                                    300
Vandenboschia_speciosa_Bel4    VRLPIRPEPSPMQLS------------------ADERSHLQLKKGRLIGMVEELDGRYRQRDQMQLVI-TSFESGTGLGAAAPYTTLARQAMSKRFRS
Vandenboschia_speciosa_Bel6    TTSMPSAASGPVQYS-----------SNVSELSQQGRQEFQMKKAKLLAMLDEVDRRYKQYYGQMQVVV-DSFESAAGPGVAKMYTSLALQTISRHFRS
Vandenboschia_speciosa_Bel10   TTSIPSFASGPEQNS-----------STVIELTREERQELQTKKAKLIAMLDEVDRRYKQYYSQMEMVV-TSFESAAGLGAAKTYTALALQTISRHFRC
Selaginella_moellendorffii     SSAVASPGTGPATVSTVAPSTNTAESKENVSVLTPDERQEYEGKKTKLLAMLQEVDRRYKQYYDQMQVI-TSFDAVAGAGAATPYTALALQAMSRYFRC
Phoenix_dactylifera_Bel1_like  SNAQESTANSPGELS------------------PSGRQDLQNKVTKLLAMLDEVDRRYKQYYHQMQIVV-SSFDVFAGSGAAKPYTALALQTISRHFRC
Theobroma_cacao_Bel1_like      SNQQECSNNSPNELS------------------HAERQELQSKTKLSSMLDEVDRRYKQYYHQMQIVV-SSFDVIAGCGAAKPYTALALQTISRHFRC
Populus_euphratica_Bel1_like   SNLQESTN-SPSELS------------------HAERQELQNKLTKLLSMLDEVDRRYKQYYHQMQIVV-SSFDVISGCGAAKPYTALALQTISRHFRC
Vitis_vinifera_Bell_like       --GGESSAKRGAELS------------------TAQRQELQMKKAKLVNMLDEVEQRYRQYHQQMQIVV-SSFEQAAGQGSAKSYTALALQTISKQFRC
Solanum_tuberosum_Bell-like_5  --ESSSRQKNEVAV-----------------ELTTAQRQELQMKKAKLLHEVEQRYRQYHHQMQIIV-LSFEQVAGIGSAKSYTQLALHAISKQFRC
Arabidopsis_thaliana_BLH1      GGGABAAGKRPVELG------------------TAERQEIQMKKAKLSNMLHEVEQRYRQYHQQMQMVISSFEQAAGIGSAKSYTSLALKTISRQFRC
Hordeum_vulgare_vulgare        --SPSSASKEPPQLS------------------PADRFEQQRKAKLISMLDEVDRRYNHYCDQMQMVV-NFFDSVMGFGAATPYTALAQKAMSRHFRC

                                                                                                    400
Vandenboschia_speciosa_Bel4    LRDSVSDQIQKVCKTLGE-ELSNVPILNRGETPRLRVLDQRLRYQRALQQIGMLQ---QQAWRPQRGLPERSVSVLRAWLFEHFLHPYPKDSEKLMLARL
Vandenboschia_speciosa_Bel6    LRDAISSQIRVASAALNE-ENPSSPALGRGEASRLRYVDQHFRQQRALQQLGMMQ---QLAWRPQRGLPAFSVWVVVVCLFLHFLLAYPKDADKIMLARQ
Vandenboschia_speciosa_Bel10   LRDAIIGQIRVASARALGE-ENFSSPCLGRGETPRLRYLDQQLRQQRALQQLGMMQ---QHAWRPQRGLPESSVSILRAWLFEHFLHPYPKDSDKLTLARQ
Selaginella_moellendorffii     LRDAITTGQIQTTCKALGE-EDVTKSITSRPLTSRLRFIDQQIRQQRAYQQYGMLQ---QHAWRPQRGLPERSVSILRAWLFEHFLHPYPKDADKMMLARQ
Phoenix_dactylifera_Bel1_like  LRDAISGQIRATRKSLGE-QDSS-ISKGGG-ISRLRYIDVDQQLRQQRAMQQFGMMQ---QHAWRPQRGLPESSVSVLRAWLFEHFLHPYPKDSEKLMLARQ
Theobroma_cacao_Bel1_like      LRDAINGQIRVTRKNLGE-QDTSENSKGVG-ISRLRYVDQQLRQQRALQQLGMMQ---QHAWRPQRGLPESSVSILRAWLFEHFLHPYPKDSDKIMLARQ
Populus_euphratica_Bel1_like   LRDAISGQIHATRKSLGE-QETSENSKGVG-ITRLRYVDQHLRQQRALHQLGMMQ---QHAWRPQRGLPESSVSILRAWLFEHFLHPYPKDSDKIMLARQ
Vitis_vinifera_Bell_like       LKDAISAQIKATSSSLGE-EDCSGGKVEG---SRLRFVDHQLRQQRALQQLGMIQ---HNAWRPQRGLPERAVSVLRAWLFEHFLHPYPKDSDKHMLAKQ
Solanum_tuberosum_Bell-like_5  LKDAIAEQVKATSKSLGE-EEGLGGKIEG---SRLKFVDHHLRQQRALQQLGMIQ---NNAWRPQRGLPERAVSVLRAWLFEHFLHPYPKDSDKIMLAKQ
Arabidopsis_thaliana_BLH1      LKEATAGQIKAANKSLGE-EDSVSGVGRFEG-SRLKFVDHHLRQQRALQQLGMIQHPSN-NAWRPQRGLPERAVSVLRAWLFEHFLHPYPKDSDKHMLAKQ
Hordeum_vulgare_vulgare        LRDAIAAQLRHTCELLGEKDAGTSSGLTKGETPRLRAIDQSLRQQRAFHHMGMME---QEAWRPQRGLPERSVSILRSWLFEHFLHPYPSDADKHLLARQ
Triticum_aestivum_Bel1_like    LKDAIAAQLRHTCELLGEKDAGTSSGLTKGETPRLRAIDQSLRQQRAFHHMGMME---QEAWRPQRGLPERSVSILRSWLFEHFLHPYPSDADKHLLARQ

                                          431
Vandenboschia_speciosa_Bel4    TGLSRSQVSNWFINARVRLWKPMVEEMYVEE
Vandenboschia_speciosa_Bel6    TGLTRNQVSNWFINARVRLWKPMVEEMYLEE
Vandenboschia_speciosa_Bel10   TGLTRNQVSNWFINARVRLWKPMVEEMYLEE
Selaginella_moellendorffii     TGLTRGQVSNWFINARVRLWKPMVEEMYQEE
Phoenix_dactylifera_Bel1_like  TGLTRSQVSNWFINARVRLWKPMIEEMYRREE
Theobroma_cacao_Bel1_like      TGLTRSQVSNWFINARVRLWKPMVEEMYKEE
Populus_euphratica_Bel1_like   TGLTRSQVSNWFINARVRLWKPMVEEMYKEE
Vitis_vinifera_Bell_like       TGLTRSQVSNWFINARVRLWKPMVEEMYLEE
Solanum_tuberosum_Bell-like_5  TGLTRSQVSNWFINARVRLWKPMVEEMYLEE
Arabidopsis_thaliana_BLH1      TGLTRSQVSNWFINARVRLWKPMVEEMYMEE
Hordeum_vulgare_vulgare        TGLSRNQVSNWFINARVRLWKPMIEEMYQQE
Triticum_aestivum_Bel1_like    TGLSRNQVSNWFINARVRLWKPMIEEMYQQE
```

Figure 2. Alignment of the deduced amino acid sequences of the VsBELL proteins from *Vandenboschia speciosa* with those of other related BELL proteins of different plant species: the lycophyte *Selaginella moellendorffii* and several seed plants. Boxes mark the different protein domains: the ZYBEL domain (green), the SKY domain (yellow), the BEL domain (blue), and the three amino acid loop extension homeodomain (purple).

3.2. Expression Analysis of TALE Homeobox Genes in V. speciosa

Reverse transcription polymerase chain reaction was performed to explore the expression levels of the selected *TALE* genes in samples from different tissues (sporophyte, spore filled sporangia, and gametophyte) belonging to the OJEN population of *V. speciosa*. We used the adenine phosphoribosyltransferase and squalene synthase genes as housekeeping genes (*HKGs*) (Table S1). The RQ values of the samples for each gene and the significance of differences (*p*-values) in *TALE* gene expression levels are indicated in Tables 2 and 3, respectively. A graphic representation of the RQ values within each population/phase is in Figure 3.

Table 2. Relative quantification (RQ) values of the transcripts ± standard error (parenthesis) for each group of samples analyzed by qPCR. The value for each sample represents the mean RQ value of five individuals taken from the same population and collecting period.

Developmental Phase	*VsKNAT1*	*VsKNAT3*	*VsKNAT4*	*VsKNAT6*	*VsBELL4*	*VsBELL6*	*VsBELL10*
Sporophyte	0.43 (0.06)	1.98 (0.35)	0.23 (0.04)	0.24 (0.03)	0.85 (0.31)	0.40 (0.06)	1.11 (0.30)
Sporangia	0.23 (0.04)	1.68 (0.40)	0.19 (0.06)	0.26 (0.10)	0.46 (0.06)	0.40 (0.04)	0.59 (0.08)
Gametophyte	0.21 (0.07)	0.21 (0.07)	1.73 (0.36)	1.46 (0.08)	0.05 (0.01)	0.31 (0.03)	0.40 (0.03)

Table 3. Estimated *p*-values of the non-parametric Mann–Whitney *U* test for the significance of differences between the relative quantification (RQ) values of the transcripts in the groups of samples. Asterisk indicates significant differences between the expression levels as detected by qPCR.

Developmental Phase	*VsKNAT1*	*VsKNAT3*	*VsKNAT4*	*VsKNAT6*	*VsBELL4*	*VsBELL6*	*VsBELL10*
Sporophyte–Gametophyte	0.076	0.009 *	0.009 *	0.009 *	0.009 *	0.175	0.028 *
Sporangia–Gametophyte	0.465	0.009 *	0.009 *	0.009 *	0.009 *	0.117	0.047 *
Sporophyte–Sporangia	0.047 *	0.465	0.601	0.917	0.602	0.917	0.251
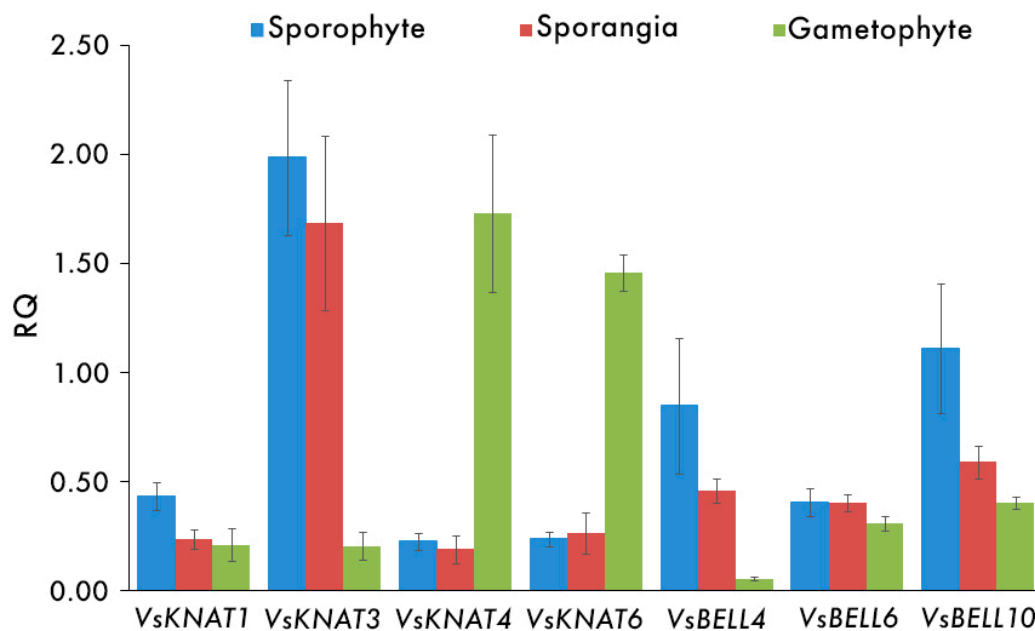
Figure 3. Graphic representation of the mean RQ values (and standard errors) of each group of samples analyzed for each gene, same as in Table 2.

While the *VsKNAT* genes 1 and 3 showed higher levels of gene expression in the sporophyte than in the gametophyte (significant only in the case of the *VsKNAT3* gene), the *VsKNAT4* and *VsKNAT6* genes showed significantly higher levels of gene expression in the gametophyte than in the sporophyte (Tables 2 and 3, and Figure 3). On the other hand, the *VsBELL* genes were always more expressed in the sporophyte than in the gametophyte. The differences were significant for the *VsBELL4* and *VsBELL10* genes. The gene expression levels of the seven genes were always slightly lower in sporangia than in sporophyte, but only in one case, that of the *VsKNAT1* gene, was this difference significant. In order to discard the possibility of variation between populations, we analyzed the expression patterns of *TALE* genes also in the VALD population. We found equivalent differences for gene expression patterns between gametophyte and sporophyte in OJEN and VALD populations, and there were not significant differences for comparisons between OJEN and VALD sporophytes on one hand, and between OJEN and VALD gametophytes, on the other (not shown).

3.3. Expression Analysis of TALE Homeobox Genes in La Almoraima, a Population Composed Only by Independent Gametophytes

In order to investigate if the *TALE* gene expression patterns of the gametophyte of ALMO were differential from the populations with both sporophytes and gametophytes, we performed qRT-PCR in ALMO and OJEN gametophyte samples. The RQ values of the samples for each gene, and the significance of differences (*p*-values) in TALE gene expression levels are indicated in Table 4. A graphic representation of the RQ values within each population is in Figure 4. There is a significant increase of

gene level expression of four genes in the ALMO population with respect to the OJEN population: *VsKNAT3*, *VsKNAT6*, *VsBELL4*, and *VsBELL6*.

Table 4. The two first lines indicate the relative quantification (RQ) values of the transcripts ± standard error (parenthesis) for each group of samples analyzed by qPCR. The value for each sample represents the mean RQ value of five individuals taken from the same population (OJEN or ALMO) and collecting period. The last line represents the estimated *p*-values of the non-parametric Mann–Whitney *U* test for the significance of differences between the relative quantification (RQ) values of the transcripts between both samples. Asterisk indicates significant differences between the expression levels as detected by qPCR.

Population	Phase	*VsKNAT1*	*VsKNAT3*	*VsKNAT4*	*VsKNAT6*	*VsBELL4*	*VsBELL6*	*VsBELL10*
OJEN	Gametophyte	0.11 (0.04)	0.09 (0.01)	0.57 (0.11)	2.58 (0.22)	0.05 (0.01)	0.31 (0.03)	0.40 (0.03)
ALMO	Gametophyte	0.19 (0.07)	0.75 (0.26)	0.79 (0.28)	4.55 (0.30)	3.14 (0.63)	0.69 (0.07)	0.44 (0.07)
OJEN–ALMO	Gametophyte	0.347	0.009 *	0.917	0.009 *	0.014 *	0.014 *	0.806
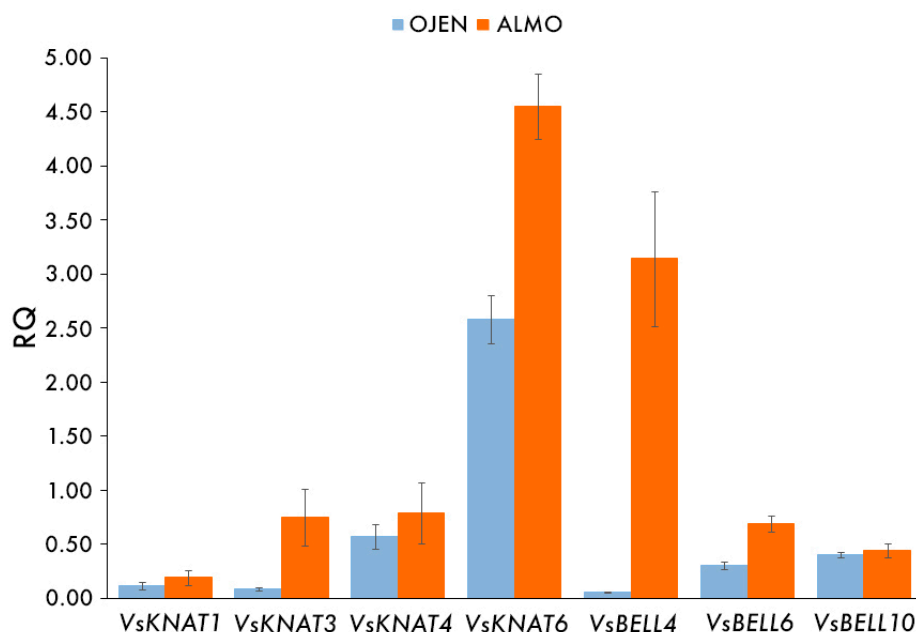
Figure 4. Graphic representation of the mean RQ values (and standard errors) of each group of samples analyzed for each gene, same as in Table 4. La Almoraima (ALMO), Canuto de Ojén Quesada (OJEN).

4. Discussion

TALE homeodomain proteins are key transcription factors controlling the gametophytic and the sporophytic developmental programs [21–25]. Ferns, as the sister lineage of seed plants, are vascular plants of great importance for the study of plant evolutionary developmental biology [1–7]. In this context, *V. speciosa* is a species with specific features, which allows the existence of an interesting model of study of fern development. Specially, as we demonstrate in this paper, the existence of independent gametophyte populations allows us to analyze changes in gene expression levels of developmental genes involved in the transition between the two phases of this fern's lifecycle. In this study, the examination of the two NGS assembled transcriptomes of *V. speciosa* revealed 22 contigs of *TALE* genes, 11 *KNOX*, and 11 *BELL* genes. The characterization and analysis of seven of these genes (four *KNOX* and three *BELL* genes) are of special interest, as it provides new information about the structure, the evolution, and the expression of *TALE* genes in ferns, a group of vascular plants for which genomic data are scarce [27,38–40]. BLAST search and sequence alignments of deduced amino acid sequences (Figures 1 and 2) demonstrate that the seven selected genes encoded TALE proteins: 2 KNOX1 proteins and 2 KNOX2 proteins, as well as 3 BELL proteins. Sequence analyses

demonstrated that these proteins are homologous to KNOX and BELL proteins of other ferns, as well as to this type of proteins in mosses, lycophytes, and seed plant.

Prior to the origin of land plants, a gene duplication in an ancestral *KNOX* gene generated the two classes of *KNOX* genes (*KNOX1* and *KNOX2*) [11,22,25–27]. The ancestral role of the *KNOX* and *BELL* genes was the regulation of sexual and zygote development as occurs in green algae and several fungi [24,25,32]. In the moss *P. patens*, both class I and class II *KNOX* genes are predominantly expressed in sporophytes, although *KNOX2* genes have also been detected in the haploid tissues, such as egg mother cells and mature eggs [24,58]. *KNOX1* genes are involved in cell proliferation during sporophyte development [25,32,58], while *KNOX2* genes are required to repress gametophyte developmental program in sporophytes, having a critical role in establishing an alternation of generations in land plants by the regulation of the gametophyte-to-sporophyte developmental transition [24,32,58]. Angiosperm class I gene function is slightly reminiscent to that of moss homologs, because both are involved in cell proliferation, whereas the neofunctionalization of class II genes was instrumental in the evolution of more complex multicellular diploid (sporophytic) generations in land plants, providing plasticity for the morphological evolution of land plant body plans [25,32]. In *Arabidopsis*, *KNOX1* activity promotes meristem maintenance and *KNOX2* activity promotes tissue differentiation [25]. In general terms, *KNOX1* genes in angiosperms are involved in the formation and maintenance of the shoot apical meristem (SAM), contributing to SAM function as well as to inflorescence and fruit development [59–62], and to flower patterning [18,22], whereas *KNOX2* genes are expressed in leaves, floral organs, seeds, and roots, having a broad tissue specificity compared to *KNOX1* genes. *KNOX2* genes confer opposing activities rather than redundant roles with *KNOX1* genes, and together, they act to direct the development of all above-ground organs of the *Arabidopsis* sporophyte [25]. The *KNOX* genes characterized to date in ferns are expressed in the sporophyte shoot apical meristems, but not in the gametophyte [27,38–40]. However, in *V. speciosa*, we find a differentiated role for each of the two genes belonging to each of the two KNOX classes. Thus, one *KNOX2* (VsKNAT3) is expressed predominantly in the sporophyte, with a residual expression level in the gametophyte, whereas one *KNOX1* (VsKNAT6) and one *KNOX2* (*VsKNAT4*) are expressed predominantly in the gametophyte, but they are also expressed in the sporophyte (at a lower level). Therefore, our analysis reveals a potentially important activity of both classes of *KNOX* genes in the maintenance and development of both the gametophyte and the sporophyte. The ALMO population is composed only of independent gametophytes that do not reproduce by sexual means, but they reproduce asexually. In this population, the expression levels of the *VsKNAT3* and *VsKNAT6* genes are increased, which supports an important role of these two genes in the control of the transition between the two phases of the lifecycle of this species.

In this context, the expression patterns in *V. speciosa* of the structurally and functionally related *BELL* genes appear relevant. These genes are expressed both in the sporophyte and in the gametophyte phases of this species. However, VsBELL4 and VsBELL10 are expressed at a higher level in the sporophyte than in the gametophyte, the expression of VsBELL4 being residual in the gametophyte, while the *VsBELL6* gene show similar expression levels in both phases. Interestingly again, as for the *VsKNAT* genes, things change in the gametophyte of the ALMO population, since the expression levels of VsBELL4 and VsBELL6 increase considerably. These data seem to also support also a relevant interactive role of *VsBELL* and *VsKNAT* genes in the alternation of generations in *V. speciosa*. The ancestral role in the life cycle progression of BELL proteins was inherited by land plants and is preserved in bryophytes [23,24,63]. In the moss *P. patens*, PpBELL1 activity is necessary for sporophyte development, it could be a master regulator for the gametophyte-to-sporophyte transition in *P. patens* and may trigger phase transition, embryogenesis, and asexual reproduction [23,63]. PpBELL1 and the rest of the *P. patens* BELL proteins can form heterodimers with each of the KNOX proteins of this species. Subsequently, KNOX–BELL complexes may induce the expression of additional *KNOX* genes. Similar TALE protein complexes function as transcriptional switches that regulate various developmental pathways in plants [18], such as SAM maintenance and inflorescence patterning in

Arabidopsis [59,64], seed germination and early seedling development in *Arabidopsis* [65], or tuber formation and root growth in potato [17,20,66].

In conclusion, this study offers novel data and opens the door for future further research on the homeodomain proteins of fern VsKNAT3 (a KNOX2 class protein), as well as VsBELL4 and VsBELL10 proteins, among others, which may control the vegetative program. On the contrary, one KNOX1 protein (VsKNAT6) and one KNOX2 protein (VsKNAT4) seem important during the gametophyte phase development. Furthermore, they might be among the key regulators of the gametophyte-to-sporophyte developmental transition in populations with alternation of generations, since some of them (VsKNAT3, VsKNAT6, VsBELL4, and VsBELL6) are upregulated in the non-alternating population, and might be triggering the vegetative propagation of the gametophyte and the repression of the sexual development. In that case, we could propose that VsKNAT4 protein may be probably involved in sexual cell development while VsKNAT6 could favor the asexual reproduction of the gametophyte. We might then further propose that *VsKNAT1*, *VsKNAT3*, *VsKNAT6*, and *VsBELL* genes could be involved in the vegetative development of both the sporophytes and the peculiar gametophytes of *V. speciosa*.

In a larger context, this paper provides new information about the structure, the evolution, and the expression of *TALE* genes in ferns, a group of vascular plants for which genomic data are scarce. *TALE* genes belong to a very ancient homeobox gene superclass that encodes for key transcription factors controlling the gametophytic and the sporophytic developmental programs, as well as the gametophyte-to-sporophyte developmental transition that characterizes plants. There is little information on the role of *TALE* genes in ferns, and this study has found new insights on the potential role of both *KNOX* and *BELL* genes in the genetic control of the development of this fern species. Thus, several of these genes were found differentially expressed both in the sporophyte and in the gametophyte phases of *V. speciosa*. Furthermore, the existence of independent gametophyte populations of this species allowed us to analyze changes in gene expression levels of the developmental genes that might be involved in the transition between the two phases of this fern's lifecycle.

References

1. Pryer, K.M.; Schneider, H.; Smith, A.R.; Cranfill, R.; Wolf, P.G.; Hunt, J.S.; Sipes, S.D. Horsetails and ferns are a monophyletic group and the closest living relatives to seed plants. *Nature* **2001**, *409*, 618–622. [CrossRef] [PubMed]

2. Pryer, K.M.; Schneider, H.; Zimmer, E.A.; Banks, J.A. Deciding among green plants for whole genome studies. *Trends Plant Sci.* **2002**, *7*, 550–554. [CrossRef]

3. Schneider, H.; Pryer, K.M.; Cranfill, R.; Smith, A.R.; Wolf, P.G. Evolution of vascular plant body plans: A phylogenetic perspective. In *Developmental Genetics and Plant Evolution*; Cronk, Q.C.B., Bateman, R.M., Harris, J.A., Eds.; Taylor and Francis: Philadelphia, PA, USA, 2002; pp. 330–364.

4. Banks, J.A.; Nishiyama, T.; Hasebe, M.; Bowman, J.L.; Gribskov, M.; Albert, V.A.; Aono, N.; Aoyama, T.; Ambrose, B.A.; Ashton, N.W.; et al. The *Selaginella* genome identifies genetic changes associated with the evolution of vascular plants. *Science* **2011**, *332*, 960–963. [CrossRef] [PubMed]

5. Schneider, H. Evolutionary morphology of ferns (monilophytes). *Annu. Plant Rev. Evol. Plant Form* **2013**, *45*, 115–140.

6. Huang, Q.; Li, W.; Fan, R.; Chang, Y. New *MADS*-box gene in fern: Cloning and expression analysis of DfMADS1 from *Dryopteris fragrans*. *PLoS ONE* **2014**, *9*, e86349. [CrossRef] [PubMed]

7. Plackett, A.R.G.; Di Stilio, V.S.; Langdale, J.A. Ferns: The missing link in shoot evolution and development. *Front. Plant Sci.* **2015**, *6*, 972. [CrossRef] [PubMed]

8. Rumsey, F.J.; Vogel, J.C.; Russell, S.J.; Barrett, J.A.; Gibby, M. Population genetics and conservation biology of the endangered fern *Trichomanes speciosum* (Hymenophyllaceae) in Scotland. *Biol. J. Linn. Soc.* **1999**, *66*, 333–344.

9. Johnson, G.N.; Rumsey, F.J.; Headley, A.D.; Sheffield, E. Adaptations to extreme low light in the fern *Trichomanes speciosum*. *New Phytol.* **2000**, *148*, 423–431. [CrossRef]

10. Makgomol, K.; Sheffield, E. Gametophyte morphology and ultrastructure of the extremely deep shade fern, *Trichomanes speciosum*. *New Phytol.* **2001**, *151*, 243–255. [CrossRef]

11. Mukherjee, K.; Brocchieri, L.; Bürglin, T.R. A comprehensive classification and evolutionary analysis of plant Homeobox genes. *Mol. Biol. Evol.* **2009**, *26*, 2775–2794. [CrossRef] [PubMed]

12. Bürglin, T.R. Homeodomain subtypes and functional diversity. In *A Handbook of Transcription Factors*; Springer: New York, NY, USA, 2011.

13. Bürglin, T.R.; Affolter, M. Homeodomain proteins: An update. *Chromosoma* **2016**, *125*, 497–521. [CrossRef] [PubMed]

14. Hamant, O.; Pautot, V. Plant development: A TALE story. *C. R. Biol.* **2010**, *333*, 371–381. [CrossRef] [PubMed]

15. Di Giacomo, E.; Iannelli, M.A.; Frugis, G. TALE and shape: How to make a leaf different. *Plants* **2013**, *2*, 317–342. [CrossRef] [PubMed]

16. Sharma, P.; Lin, T.; Grandellis, C.; Yu, M.; Hannapel, D.J. The BEL1-like family of transcription factors in potato. *J. Exp. Bot.* **2014**, *65*, 709–723. [CrossRef] [PubMed]

17. Ghate, T.H.; Sharma, P.; Kondhare, K.R.; Hannapel, D.J.; Banerjee, A.K. The mobile RNAs, *StBEL11* and *StBEL29*, suppress growth of tubers in potato. *Plant Mol. Biol.* **2017**, *93*, 563–578. [CrossRef] [PubMed]

18. Hay, A.; Tsiantis, M. A KNOX family TALE. *Curr. Opin. Plant Biol.* **2009**, *12*, 593–598. [CrossRef] [PubMed]

19. Hannapel, D.J. A model system of development regulated by the long-distance transport of mRNA. *J. Integr. Plant Biol.* **2010**, *52*, 40–52. [CrossRef] [PubMed]

20. Lin, T.; Sharma, P.; Gonzalez, D.H.; Viola, I.L.; Hannapel, D.J. The impact of the long-distance transport of a BEL1-like messenger RNA on development. *Plant Physiol.* **2013**, *161*, 760–772. [CrossRef] [PubMed]

21. Liu, Y.; You, S.; Taylor-Teeples, M.; Li, W.L.; Schuetz, M.; Brady, S.M.; Douglas, C.J. BEL1-LIKE HOMEODOMAIN6 and KNOTTED ARABIDOPSIS THALIANA7 interact and regulate secondary cell wall formation via repression of *REVOLUTA*. *Plant Cell* **2014**, *26*, 4843–4861. [CrossRef] [PubMed]

22. Gao, J.; Yang, X.; Zhao, W.; Lang, T.; Samuelsson, T. Evolution, diversification, and expression of KNOX proteins in plants. *Front. Plant Sci.* **2015**, *6*, 882. [CrossRef] [PubMed]

23. Horst, N.A.; Katz, A.; Pereman, I.; Decker, E.L.; Ohad, N.; Reski, R. A single homeobox gene triggers phase transition, embryogenesis and asexual reproduction. *Nat. Plants* **2016**, *2*. [CrossRef] [PubMed]

24. Sakakibara, K.; Ando, S.; Yi, H.K.; Tamada, Y.; Hiwatashi, Y.; Murata, T.; Deguchi, H.; Hasebe, M.; Bowman, J.L. *KNOX2* genes regulate the haploid-to-diploid morphological transition in land plants. *Science* **2013**, *339*, 1067–1070. [CrossRef] [PubMed]

25. Furumizu, C.; Alvarez, J.P.; Sakakibara, K.; Bowman, J.L. Antagonistic roles for *KNOX1* and *KNOX2* genes in patterning the land plant body plan following an ancient gene duplication. *PLoS Genet.* **2015**, *11*, e1004980. [CrossRef] [PubMed]

26. Kerstetter, R.; Vollbrecht, E.; Lowe, B.; Veit, B.; Yamaguchi, J.; Hake, S. Sequence-Analysis and Expression Patterns Divide the Maize Knotted1-Like Homeobox Genes into 2 Classes. *Plant Cell* **1994**, *6*, 1877–1887. [CrossRef] [PubMed]

27. Bharathan, G.; Janssen, B.J.; Kellogg, E.A.; Sinha, N. Did homeodomain proteins duplicate before the origin of angiosperms, fungi, and metazoa? *Proc. Natl. Acad. Sci. USA* **1997**, *94*, 13749–13753. [CrossRef] [PubMed]

28. Raven, J.A. Comparative physiology of plant and arthropod land adaptation. *Philos. Trans. R. Soc. Lond. B* **1985**, *309*, 273–288. [CrossRef]

29. Raven, J.A. Plant responses to high O_2 concentrations: Relevance to previous high O_2 episode. *Palaeogeogr. Palaeoclimatol. Palaeoecol.* **1991**, *97*, 19–38. [CrossRef]

30. Edwards, D.; Fanning, U.; Richardson, J.B. Stomata and sterome in early land plants. *Nature* **1986**, *323*, 438–440. [CrossRef]

31. Sztein, A.E.; Cohen, J.D.; Slovin, J.P.; Cooke, T.J. Auxin metabolism in representative land plants. *Am. J. Bot.* **1995**, *82*, 1514–1521. [CrossRef]

32. Tsuda, K.; Hake, S. Diverse functions of KNOX transcription factors in the diploid body plan of plants. *Curr. Opin. Plant Biol.* **2015**, *27*, 91–96. [CrossRef] [PubMed]

33. Zimmer, A.D.; Lang, D.; Buchta, K.; Rombauts, S.; Nishiyama, T.; Hasebe, M.; Van de Peer, Y.; Rensing, S.A.; Reski, R. Reannotation and extended community resources for the genome of the non-seed plant *Physcomitrella patens* provide insights into the evolution of plant gene structures and functions. *BMC Genom.* **2013**, *14*, 498. [CrossRef] [PubMed]

34. Salmi, M.L.; Bushart, T.J.; Stout, S.C.; Roux, S.J. Profile and analysis of gene expression changes during early development in germinating spores of *Ceratopteris richardii*. *Plant Physiol.* **2005**, *138*, 1734–1745. [CrossRef] [PubMed]

35. Matasci, N.; Hung, L.H.; Yan, Z.; Carpenter, E.J.; Wickett, N.J.; Mirarab, S.; Nguyen, N.; Warnow, T.; Ayyampalayam, S.; Barker, M.; et al. Data access for the 1000 Plants (1KP) project. *GigaScience* **2014**, *3*, 17. [CrossRef] [PubMed]

36. Wolf, P.G.; Sessa, E.B.; Marchant, D.B.; Li, F.W.; Rothfels, C.J.; Sigel, E.M.; Gitzendanner, M.A.; Visger, C.J.; Banks, J.A.; Soltis, D.E.; et al. An exploration into fern genome space. *Genome Biol. Evol.* **2015**, *7*, 2533–2544. [CrossRef] [PubMed]

37. Frank, M.H.; Edwards, M.B.; Schultz, E.R.; McKain, M.R.; Fei, Z.; Sørensen, I.; Rose, J.K.C.; Scanlon, M.J. Dissecting the molecular signatures of apical cell-type shoot meristems from two ancient land plant lineages. *New Phytol.* **2015**, *207*, 893–904. [CrossRef] [PubMed]

38. Harrison, C.J.; Corley, S.B.; Moylan, E.C.; Alexander, D.L.; Scotland, R.W.; Langdale, J.A. Independent recruitment of a conserved developmental mechanism during leaf evolution. *Nature* **2005**, *434*, 509–514. [CrossRef] [PubMed]

39. Sano, R.; Juárez, C.M.; Hass, B.; Sakakibara, K.; Ito, M.; Banks, J.A.; Hasebe, M. *KNOX* homeobox genes potentially have similar function in both diploid unicellular and multicellular meristems, but not in haploid meristems. *Evol. Dev.* **2005**, *7*, 69–78. [CrossRef] [PubMed]

40. Ambrose, B.A.; Vasco, A. Bringing the multicellular fern meristem into focus. *New Phytol.* **2016**, *210*, 790–793. [CrossRef] [PubMed]

41. FastQC. A quality control tool for high throughput sequence data. Available online: https://www.bioinformatics.babraham.ac.uk/projects/fastqc/ (accessed on 16 October 2017).

42. Simpson, J.T.; Wong, K.; Jackman, S.D.; Schein, J.E.; Jones, S.J.; Birol, I. ABySS: A parallel assembler for short read sequence data. *Genome Res.* **2009**, *19*, 1117–1123. [CrossRef] [PubMed]

43. Robertson, G.; Schein, J.; Chiu, R.; Corbett, R.; Field, M.; Jackman, S.D.; Mungall, K.; Lee, S.; Okada, H.M.; Qian, J.Q.; et al. De novo assembly and analysis of RNA-seq data. *Nat. Methods* **2010**, *7*, 909–912. [CrossRef] [PubMed]

44. Huang, X.; Madan, A. CAP3: A DNA Sequence Assembly Program. *Genome Res.* **1999**, *9*, 868–877. [CrossRef] [PubMed]

45. Altschul, S.F.; Madden, T.L.; Schäffer, A.A.; Zhang, J.; Zhang, Z.; Miller, W.; Lipman, D.J. Gapped BLAST and PSI-BLAST: A new generation of protein database search programs. *Nucleic Acids Res.* **1997**, *25*, 3389–3402. [CrossRef] [PubMed]

46. Conesa, A.; Götz, S.; García-Gómez, J.M.; Terol, J.; Talón, M.; Robles, M. Blast2GO: A universal tool for annotation, visualization and analysis in functional genomics research. *Bioinformatics* **2005**, *21*, 3674–3676. [CrossRef] [PubMed]

47. Li, H.; Durbin, R. Fast and accurate short read alignment with Burrows-Wheeler transform. *Bioinformatics* **2009**, *25*, 1754–1760. [CrossRef] [PubMed]

48. Anders, S.; Pyl, P.T.; Huber, W. HTSeq—A Python framework to work with high-throughput sequencing data. *Bioinformatics* **2014**, *31*, 166–169. [CrossRef] [PubMed]

49. Benjamini, Y.; Hochberg, Y. Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing. *J. R. Stat. Soc. Ser. B.* **1995**, *57*, 289–300.

50. Hall, T.A. BioEdit: A user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. *Nucleic Acids Symp. Ser.* **1999**, *41*, 95–98.

51. Higgins, D.G.; Sharp, P.M. CLUSTAL: A package for performing multiple sequence alignment on a microcomputer. *Gene* **1988**, *73*, 237–244. [CrossRef]

52. Le Bail, A.; Scholz, S.; Kost, B. Evaluation of Reference Genes for RT qPCR analyses of structure-specific and hormone regulated gene expression in *Physcomitrella patens* Gametophytes. *PLoS ONE* **2013**, *8*, e70998. [CrossRef] [PubMed]

53. Rozen, S.; Skaletsky, H. Primer3 on the WWW for general users and for biologist programmers. *Methods Mol. Biol.* **2000**, *132*, 365–386. [PubMed]

54. Vandesompele, J.; De Preter, K.; Pattyn, F.; Poppe, B.; Van Roy, N.; De Paepe, A.; Speleman, F. Accurate normalization of real-time quantitative RT-PCR data by geometric averaging of multiple internal control genes. *Genome Biol.* **2002**, *3*. [CrossRef]

55. TIBCO Statistica. Available online: http://statistica.io/products/ (accessed on 17 October 2017).

56. GeneBank. Available online: https://www.ncbi.nlm.nih.gov/genbank/ (accessed on 17 October 2017).

57. The European Nucleotide Archive (ENA). Available online: https://www.ebi.ac.uk/ena (accessed on 17 October 2017).

58. Sakakibara, K.; Nishiyama, T.; Deguchi, H.; Hasebe, M. Class 1 *KNOX* genes are not involved in shoot development in the moss *Physcomitrella patens* but do function in sporophyte development. *Evol. Dev.* **2008**, *10*, 555–566. [CrossRef] [PubMed]

59. Byrne, M.E.; Simorowski, J.; Martienssen, R.A. ASYMMETRIC LEAVES1 reveals *KNOX* gene redundancy in *Arabidopsis*. *Development* **2002**, *129*, 1957–1965. [PubMed]

60. Douglas, S.J.; Chuck, G.; Dengler, R.E.; Pelecanda, L.; Riggs, C.D. KNAT1 and ERECTA regulate inflorescence architecture in *Arabidopsis*. *Plant Cell* **2002**, *14*, 547–558. [CrossRef] [PubMed]

61. Venglat, S.P.; Dumonceaux, T.; Rozwadowski, K.; Parnell, L.; Babic, V.; Keller, W.; Martienssen, R.; Selvaraj, G.; Datla, R. The homeobox gene *BREVIPEDICELLUS* is a key regulator of inflorescence architecture in *Arabidopsis*. *Proc. Natl. Acad. Sci. USA* **2002**, *99*, 4730–4735. [CrossRef] [PubMed]

62. Ragni, L.; Belles-Boix, E.; Günl, M.; Pautot, V. Interaction of *KNAT6* and *KNAT2* with *BREVIPEDICELLUS* and *PENNYWISE* in *Arabidopsis* inflorescences. *Plant Cell* **2008**, *20*, 888–900. [CrossRef] [PubMed]

63. Harrison, C.J. Development and genetics in the evolution of land plant body plans. *Philos. Trans. R. Soc. B* **2017**, *372*. [CrossRef] [PubMed]

64. Roeder, A.H.; Ferrandiz, C.; Yanofsky, M.F. The role of the REPLUMLESS homeodomain protein in patterning the *Arabidopsis* fruit. *Curr. Biol.* **2003**, *13*, 1630–1635. [CrossRef] [PubMed]

65. Kim, D.; Cho, Y.H.; Ryu, H.; Kim, Y.; Kim, T.H.; Hwang, I. BLH1 and KNAT3 modulate ABA responses during germination and early seedling development in *Arabidopsis*. *Plant J.* **2013**, *75*, 755–766. [CrossRef] [PubMed]

66. Chen, H.; Rosin, F.M.; Prat, S.; Hannapel, D.J. Interacting transcription factors from the three-amino acid loop extension superclass regulate tuber formation. *Plant Physiol.* **2003**, *132*, 1391–1404. [CrossRef] [PubMed]