

Article

Copaifera langsdorffii Novel Putative Long Non-Coding RNAs: Interspecies Conservation Analysis in Adaptive Response to Different Biomes

Monica F. Danilevicz¹, Kanhu C. Moharana² , Thiago M. Venancio², Luciana O. Franco³, Sérgio R. S. Cardoso³, Mônica Cardoso³, Flávia Thiebaut¹ , Adriana S. Hemerly¹, Francisco Prosdocimi⁴  and Paulo C. G. Ferreira^{1,*} 

¹ Laboratório de Biologia Molecular de Plantas, Instituto de Bioquímica Médica Leopoldo de Meis, Universidade Federal do Rio de Janeiro, Rio de Janeiro 21941-599, Brazil; monica.danilevicz@gmail.com (M.F.D.); flaviabqi@gmail.com (F.T.); hemerly@bioqmed.ufrj.br (A.S.H.)

² Laboratório de Química e Função de Proteínas e Peptídeos, Centro de Biotecnologia, Universidade Estadual do Norte Fluminense Darcy Ribeiro, Rio de Janeiro 28013-602, Brazil; kcm.eid@gmail.com (K.C.M.); thiago.venancio@gmail.com (T.M.V.)

³ Instituto de Pesquisas Jardim Botânico do Rio de Janeiro, Diretoria de Pesquisa Científica, Rio de Janeiro 22460-030, Brazil; lfranco@jbrj.gov.br (L.O.F.); sergio@jbrj.gov.br (S.R.S.C.); mcardoso@jbrj.gov.br (M.C.)

⁴ Laboratório de Genômica e Biodiversidade, Instituto de Bioquímica Médica Leopoldo de Meis, Universidade Federal do Rio de Janeiro, Rio de Janeiro 21941-599, Brazil; prosdocimi@bioqmed.ufrj.br

* Correspondence: paulof@bioqmed.ufrj.br; Tel.: +55-2198884-2585

Received: 17 July 2018; Accepted: 28 September 2018; Published: 8 October 2018



Abstract: Long non-coding RNAs (lncRNAs) are involved in multiple regulatory pathways and its versatile form of action has disclosed a new layer in gene regulation. lncRNAs have their expression levels modulated during plant development, and in response to stresses with tissue-specific functions. In this study, we analyzed lncRNA from leaf samples collected from the legume *Copaifera langsdorffii* Desf. (copaíba) present in two divergent ecosystems: Cerrado (CER; Ecological Station of Botanical Garden in Brasília, Brazil) and Atlantic Rain Forest (ARF; Rio de Janeiro, Brazil). We identified 8020 novel lncRNAs, and they were compared to seven Fabaceae genomes and transcriptomes, to which 1747 and 2194 copaíba lncRNAs were mapped, respectively, to at least one species. The secondary structures of the lncRNAs that were conserved and differentially expressed between the populations were predicted using in silico methods. A few selected lncRNA were confirmed by RT-qPCR in the samples from both biomes; Additionally, the analysis of the lncRNA sequences predicted that some might act as microRNA (miRNA) targets or decoys. The emerging studies involving lncRNAs function and conservation have shown their involvement in several types of biotic and abiotic stresses. Thus, the conservation of lncRNAs among Fabaceae species considering their rapid turnover, suggests they are likely to have been under functional conservation pressure. Our results indicate the potential involvement of lncRNAs in the adaptation of *C. langsdorffii* in two different biomes.

Keywords: novel lncRNA; lncRNA conservation; *Copaifera*; epigenetics; adaptive response

1. Introduction

Copaifera is a genus of native trees from Latin American tropical regions and Western Africa. The species *Copaifera langsdorffii*, *Copaifera reticulata*, *Copaifera cearensis*, *Copaifera multijuga* among others are popularly known as “copaíba” [1,2]. The *Copaifera* sp. oil is extracted through V-shaped cut in the stem bark and has been used by indigenous populations from the Brazilian Amazon as a powerful antimicrobial, anti-inflammatory and for “overall healing purposes” [3]. Many of these

alleged features have been investigated by pharmacological studies, confirming its anti-inflammatory capacity [3–5] and proving it to be an efficient alternative to treat dental infections [6], gastrointestinal disorders [7], endometriosis [8], skin ulcers [9,10], and to be applied to skin scaffold implants increasing tissue angiogenesis [11]. The copaiba oil resin has been shown to exert larvicidal activity against *Aedes aegypti* [12,13], and antibacterial activity in vitro [6,14–16]. The leaf extracts also presented leishmanicidal and antimalarial activities [17–21], also being effective as a biopesticide against lepidoptera [22].

The amount of copaiba oil production is influenced by the climate and soil conditions, as more oil is produced in locations with clay soil, during the wet season [23]. In Brazil, *C. langsdorffii* Desf. is a widely occurring species, included in a broad range of ecosystems like Cerrado, Atlantic Rain Forest and Caatinga, which are very distinct biomes [23,24], requiring diverse adaptive mechanisms. Such plasticity to adapt to different biomes is a complex regulation, involving several genetic, evolutionary and epigenetic fine tuning, which may also include long non-coding RNAs (lncRNAs).

lncRNAs are considered to be RNA transcripts, defined as longer than 200 bp, with no apparent protein coding capacity [25–30]. They hold many resemblances with messenger RNA (mRNA), such as similar epigenetic marks to promote expression and binding sites for RNA polymerase II at their genome loci. They are often polyadenylated at 3' end and receive 5' CAPs, especially when acting outside the nucleus [25–27,30,31]. Because they are (i) less conserved than mRNA at the sequence level across different species and (ii) regularly transcribed at low levels, lncRNAs were once considered transcriptional noise. lncRNAs tends to form secondary and tertiary structures, in which the molecule conformation is crucial to regulate their targets, leading to their functional domains and genomic positions to be more conserved than their sequence [32].

lncRNAs interspecies conservation may be influenced by the regulatory mechanism they play, whereas their sequence conservation depends on whether (a) the lncRNA molecule acts as a regulator or (b) the simple transcription of the lncRNA regulates its target gene [25]. Most functional lncRNAs undergo post-transcriptional processing and retain higher conservation of splice sites. This indicates that they most likely function in the mature form [25,30]. In vertebrates, it is argued that transposable elements and bidirectional transcription may play an important role in the evolution and rapid turnover of lncRNAs [26,29].

There are several types of lncRNAs classified according to their genomic positions such as: Sense, natural antisense, bidirectional, intronic and intergenic; and they may also act through cis or trans regulation [28,31,33]. lncRNAs exhibit relatively low expression patterns compared to mRNAs, showing a specific profile depending on the (i) tissue or cell type observed [34,35], (ii) developmental stage [36–38] and (iii) environmental stress response [37,39–41]

A considerable amount of lncRNAs may act as chromatin regulators [42–45]. For example, *APOLO*, a lncRNA responsive to auxin, interacts with the chromatin, leading to a loop formation encompassing the *PID* gene (key regulator of polar auxin transport) and regulating its expression [46]. *HOTAIR*, *ANRIL* and *KCNQ1OT1* are also known to bind to more than one histone-modifying complex, acting as regulators [47,48]. *COLD AIR* [49], *COOL AIR* [50] and *COLD WRAP* [51] are lncRNA described to act regulating the FLOWERING LOCUS C (FLC), associated with the Polycomb protein complex to stably repress FLC during the vernalization process, prompting adaptive fitness and development in *Arabidopsis thaliana*. In rice, it was shown that photoperiodic-sensitive male sterility (PSMS) is carried out by a lncRNA called *LDMAR*, in which a single-nucleotide polymorphism (SNP) led to a change in its secondary structure and subsequently the repression of its expression at long day conditions [52]. Several other lncRNA have their functions experimentally tested, such as *HDI* promoting photomorphogenesis in red light [53], and *NERDL* association to wood formation in *Populus tormentosa* [54].

In order to modulate the plants response to adaptive stress, such as inorganic phosphorus starvation, the lncRNA *cis-NAT PHO1;2* acts regulating phosphate (Pi) homeostasis as a translational enhancer of *OsPHO1;2*, increasing its uptake when over-expressed in a mutated rice lineage [55].

Evidence of lncRNA involved in stress response and adaptation was also observed in nitrogen deficient *Populus* [40], maize under drought and nutrient stress [56,57] and *Arabidopsis thaliana* submitted to salt stress [58]. Moreover, lncRNAs have also been found to act as small RNA (sRNA) precursors [27,59] or acting as bait to sRNA regulation by target mimicry [60–62]. Collectively, lncRNAs have been implicated in various cell and molecular processes, including post-transcriptional regulation, post-translational regulation of protein activity and protein re-localization, organization of protein complexes, cell–cell signaling and intrinsically connected to adaptive fitness and overall plant processes [28,29,31,33].

In this study, we identified 8,020 novel lncRNAs through bioinformatics analysis, of which approximately 565 were shown to be up-regulated above 5 times in Cerrado (CER) and Atlantic Rain Forest (ARF) populations, which might be related to their adaptation to such diverse environments. lncRNAs tend to have their primary sequence evolving under a relaxed constraint, being less likely to present high interspecies sequence similarity. Nevertheless, through the comparison to other Fabaceae genomes, there were 1747 putative lncRNA conserved, some of them presenting different expression profiles among the plant populations. Additionally, we aligned the copaiba transcripts to same family species transcriptomes and found 2194 aligned to at least one species transcriptome. There were 1141 lncRNAs which overlapped, being aligned in both analyses, from which some of them had differential expression. Also, we were able to identify one single transcript that matched a known *Glycine max* lncRNA [63]. Amongst the conserved lncRNAs with higher degrees of differential expression that had their secondary structure stability predicted, 186 transcripts are predicted to be stable. The emerging studies involving lncRNAs function and conservation have shown their involvement to several types of biotic and abiotic stress. Thus, the conservation of lncRNAs among Fabaceae species considering their rapid turnover, suggests they are likely to have been under functional conservation pressure [32,64,65]. The differential expression observed for hundreds of lncRNAs suggests that they take part in regulatory pathways that lead to adaptive responses in copaiba, which has high health and economic interest.

2. Results

2.1. Identification of Novel and Differentially Expressed lncRNA

In order to identify lncRNAs and elucidate their adaptive roles, we sequenced the transcriptomes of *C. langsdorffii* leaves obtained from trees growing in two different ecosystems, the (i) Atlantic Rain Forest (ARF, humid condition) and the (ii) Cerrado (CER, dry forest) in Brazil during the drought season for both locations. Approximately 75 million high quality reads from each library were used for *de novo* assembly, and there was a total of 138,175 and 199,556 transcripts assembled from ARF and CER, respectively. A series of filters were applied to these transcripts to remove potentially coding transcripts, such as coding potential calculator (CPC) (v0.9-r2) [66] and predictor of long non-coding RNAs and messenger RNAs based on an improved k-mer scheme (PLEK) (v1.2) [67], and only transcripts classified as non-coding by both software were compared to establish one-to-one correspondence between CER and ARF. Reads were remapped against the assembled transcriptomes with Bowtie2 [68] and transcriptional levels estimated with Cufflinks v2.2.1 [69] (Table 1). Transcripts with at least 1 RPKM (reads per kilobase of transcript per million mapped reads) were kept for downstream analyses (8020 transcripts) (Table S1). The majority of these (2312 transcripts) presented over 2-fold regulation, while 565 transcripts were regulated above 5-fold (Figure 1, Table S2).

Table 1. De novo transcript assembly and lncRNA identification pipeline results of copaiba from Atlantic Rain Forest (ARF) and Cerrado (CER).

Sample	Total Transcripts	Longest Contig per Cluster	CPC lncRNA Prediction	PLEK lncRNA Prediction	Overlapped lncRNA Predicted	One-to-One Correspondence (RPKM > 1)
ARF	138,175	94,815	67,251	86,608	64,801	8020
CER	199,556	140,011	102,804	129,710	99,570	8020

Abbreviations: lncRNA: long non-coding RNA; CPC: coding potential calculator; PLEK: predictor of long non-coding RNAs and messenger RNAs based on an improved k-mer scheme; RPKM: reads per kilobase of transcript per million mapped reads.

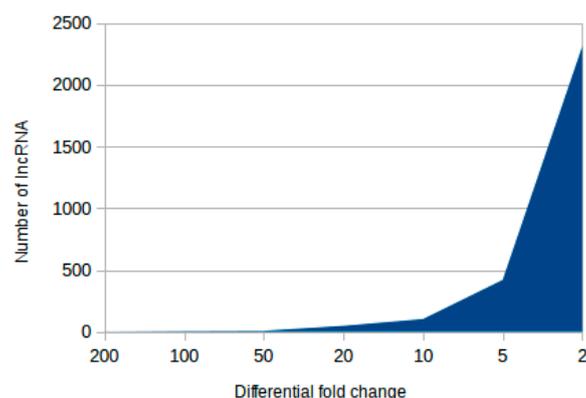


Figure 1. Putative lncRNAs identified in CER and ARF samples and its fold change regulation in comparison to each other: There were 2893 differentially regulated transcripts identified from a total of 8020 copaiba lncRNAs. The majority of the transcripts were 2 to 5 times differently expressed on either sample, yet there were 565 transcripts regulated above 5-fold on either sample.

2.2. Interspecies lncRNA Conservation Analysis

2.2.1. Positional Conservation and Genome Alignment Analysis

Since there is no reference genome for *C. langsdorffii*, we performed a similarity search using Bowtie2 (v2.3.4.1) [68] against the genomes of the following related species: (i) *Vicia faba*, (ii) *Glycine max* [70], (iii) *Medicago truncatula* [71], (iv) *Phaseolus vulgaris* [72], (v) *Lotus japonica* [73], (vi) *Vigna unguiculata* [74] and (vii) *Cicer reticulatum* [75]. From the alignments, we found 1747 and 1879 transcripts respectively from ARF and CER samples to match at least one of the genomes used. From which we noticed 156 transcripts aligned to all seven Fabaceae genomes in both CER and ARF samples (Figure 2), then being considered highly sequence conserved lncRNA transcripts [29,64,65].

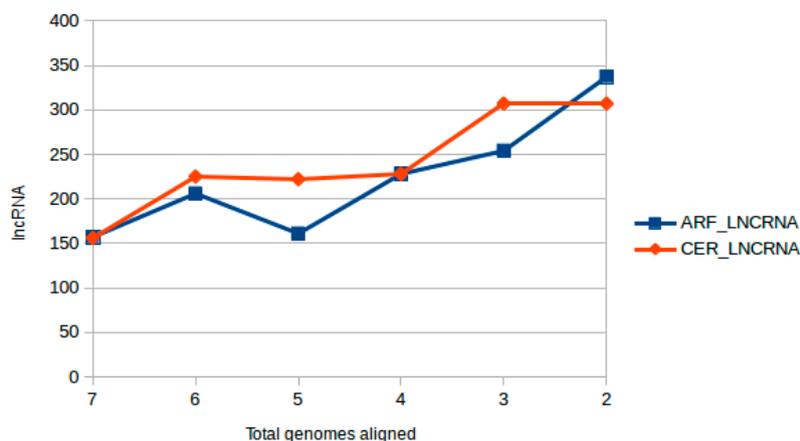


Figure 2. lncRNA interspecies conservation and genome alignment analysis: the graph shows the total number of genomes versus the total number of transcripts aligned to Fabaceae genomes (*Vicia faba*, *Glycine max*, *Medicago truncatula*, *Phaseolus vulgaris*, *Lotus japonica*, *Vigna unguiculata* and *Cicer reticulatum*). The blue line represents the lncRNA gathered from the ARF samples and the red line represents the lncRNA gathered from the CER sample, their alignment profile is very similar as expected, around 1800 transcripts aligned to at least one of the genomes used.

From the subset of 156 copaiba lncRNAs aligned to all seven Fabaceae genomes, 45 transcripts presented above 1-fold differential regulation (Figure 3). To understand if the copaiba lncRNA transcripts aligned to all seven Fabaceae genomes hold positional conservation, we selected 10 transcripts exhibiting at least 5-fold up-regulation to retrieve the information regarding their genomic locations. For this analysis, we used the reference genomes of *P. vulgaris*, *G. max* and *M. truncatula*, as they are better assembled and annotated. However, the genes closely located to the lncRNA loci are described only as “hypothetical protein-coding”, or “plant-like protein”. Furthermore, none of the 6133 lncRNA annotated in the *G. max* genome were located close to copaiba putative lncRNA alignment locations. The *M. truncatula* or *P. vulgaris* annotation files had no information on lncRNAs.

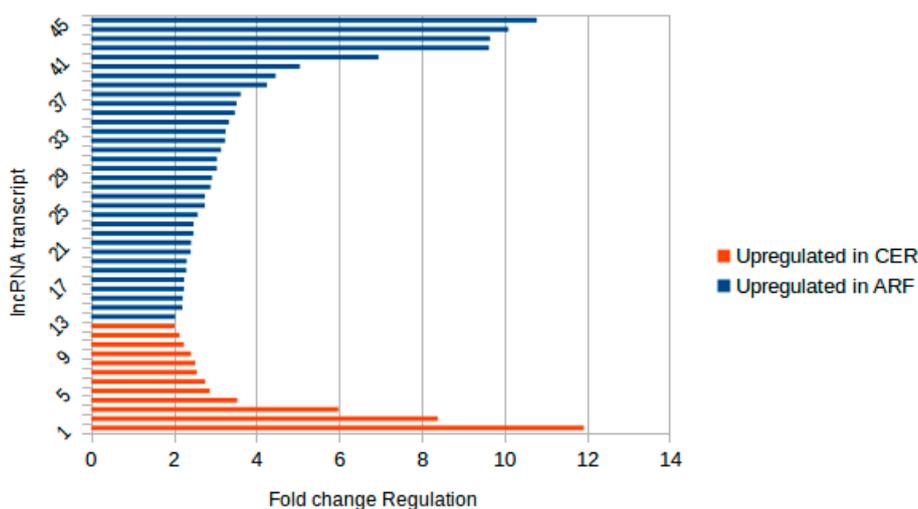


Figure 3. Subset of transcripts which aligned to all Fabaceae genomes presenting differential expression: From 156 transcripts which aligned to all genomes analyzed, there were 45 copaiba lncRNAs upregulated in either condition represented in this graph. Each transcript is represented by a single bar. In red are indicated the lncRNAs upregulated in CER samples in relation to ARF. In blue are indicated the lncRNAs upregulated in ARF samples in relation to CER. The x axis indicates the fold change regulation of the transcripts.

2.2.2. Identification of Putative *C. langsdorffii* lncRNAs in EST Sequences of Other Fabaceae Species

We downloaded expressed sequence tag (EST) and complementary DNA (cDNA) libraries from Phytozome and NCBI databases from six Fabaceae species to identify whether the copaiba lncRNA transcripts that aligned to the genomes could also be found in available transcriptome libraries, which comprise both mRNA and poly-A non-coding RNA (ncRNA). lncRNAs have a specific expression profile, considered to be lower than mRNAs. The likelihood of expression is highly influenced by tissue and specific condition [25,37,39]; therefore, the identification of copaiba lncRNA in Fabaceae transcriptome is susceptible to being underestimated. This is due to the fact that transcripts specific profile of expression and relatively low amounts may not be observed, even if present in the reference transcriptomes.

We used Basic Local Alignment Search Tool (BLASTN) (v2.2.31+) [76] to search putative copaiba lncRNAs against downloaded transcript sequences. Two identity thresholds, 50% and 90%, were used in this analysis, along with e-value and coverage threshold of 10^{-15} and 50%, respectively. We found that 27% (2194 lncRNA) aligned to at least one species' transcriptome (using 50% identity filter on BLASTN) (Figure 4). While only 3.3% (264 lncRNA) aligned with 90% identity to at least one Fabaceae expressed transcripts library and a single transcript (JCF44_0000056614) aligned to five species. It is remarkable to notice that 227 transcripts aligned to *G. max* with more than 90% identity, which might be due to *G. max* to being more intensively studied species than the others (Figure 5).

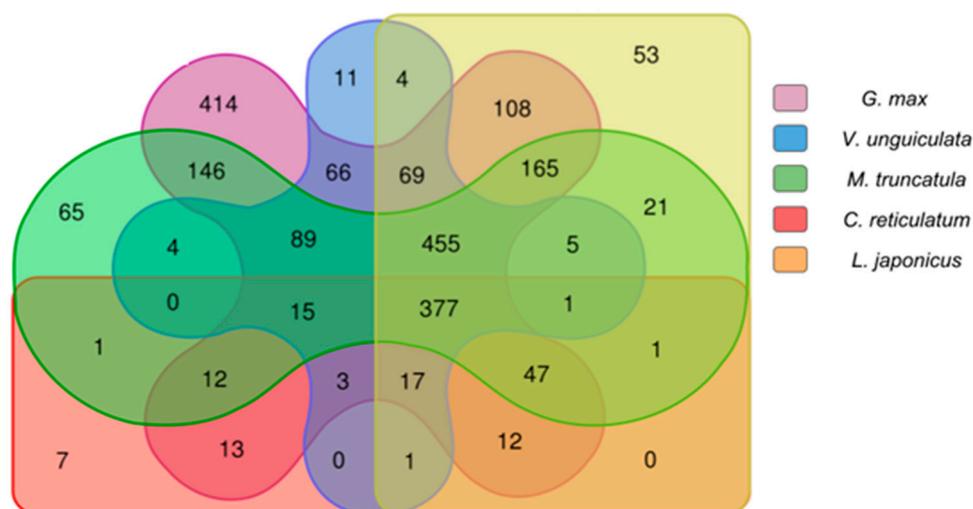


Figure 4. Count of copaiba lncRNAs aligned to each Fabaceae transcriptome with BLASTN (50% identity): This diagram shows the amount of copaiba lncRNA aligned to the transcriptome of each species, segregated by a color pattern indicated on the legend. In the colored overlapped area are the transcripts which were aligned to more than one species transcriptome, and its respective amount. There are only five species shown in this diagram, for illustration purposes we left out *P. vulgaris*.

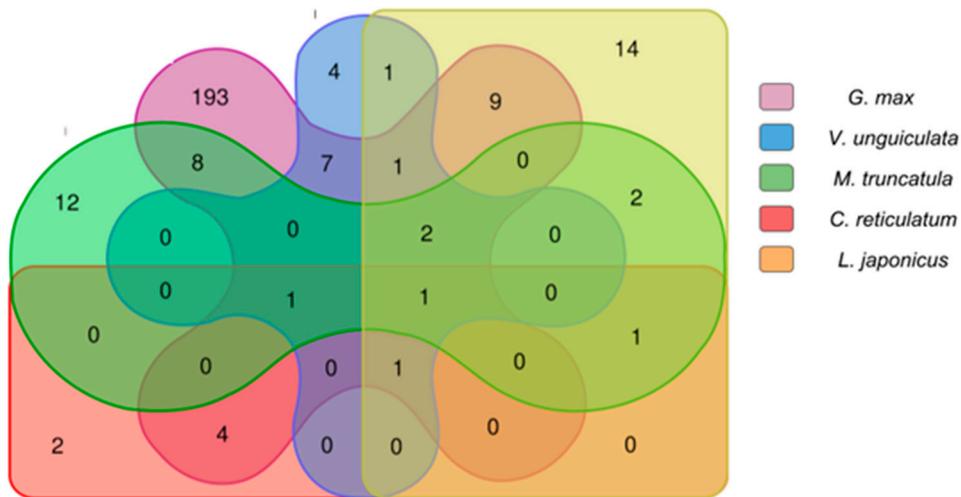


Figure 5. Count of copalba lncRNAs aligned to each Fabaceae transcriptome with BLASTN (90% identity): This diagram shows the amount of copalba lncRNA aligned to the transcriptome of each species, segregated by a color pattern indicated on the legend. In the colored overlapped area are the transcripts which were aligned to more than one species transcriptome, and its respective amount. There are only five species shown in this diagram, for illustration purposes we left out *P. vulgaris*.

2.2.3. Comparison Analysis of Conserved Putative lncRNAs

To assess whether the copalba lncRNAs aligned to Fabaceae transcribed libraries were also aligned to Fabaceae genomes, we used the Bowtie2 [68] aforementioned result, aiming to determine the number of transcripts that are present in the genome and that also are actively transcribed in the cDNA, EST libraries used (Figure 6, Table S3). Thus, we identified 1141 transcripts that aligned to both genome and transcriptome comparisons of at least one species. From this total, there were 36 copalba lncRNA transcripts which aligned to both genome and transcriptome of six Fabaceae species (Figure 7).

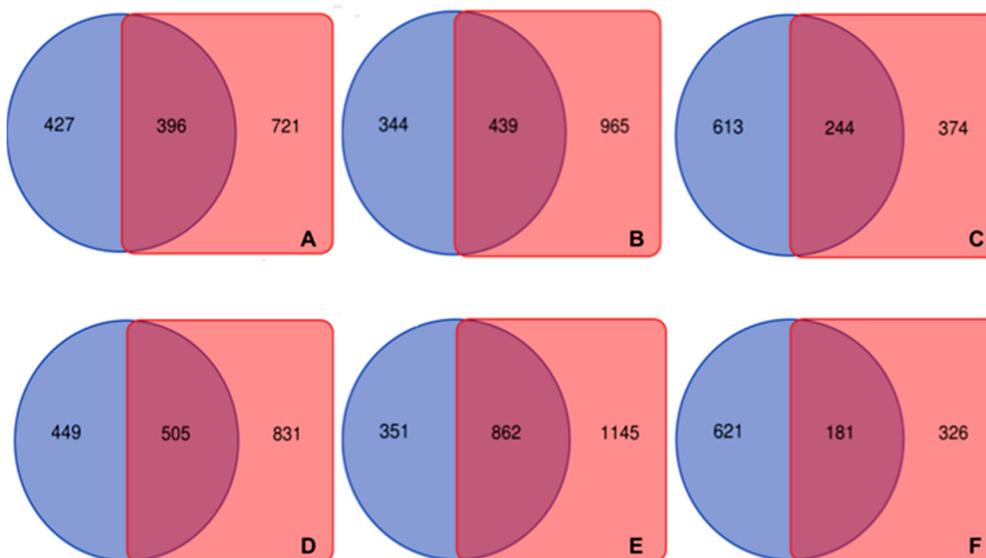


Figure 6. BLASTN comparison of copalba lncRNA aligned to each species genome and complementary DNA (cDNA): In the venn diagram is compared the amount of copalba transcripts aligned to the genome (blue circle), to the transcriptome (red square) or to both (overlapped area). (A) is the *V. unguiculata* comparison, (B) is *M. truncatula*, (C) is *P. vulgaris*, (D) is *L. japonicus*, (E) is *G. max* and (F) is *C. reticulatum*. It is possible to observe that *G. max*, *L japonicus* and *M. truncatula* presented a higher number of overall aligned transcripts and also overlapped ones.

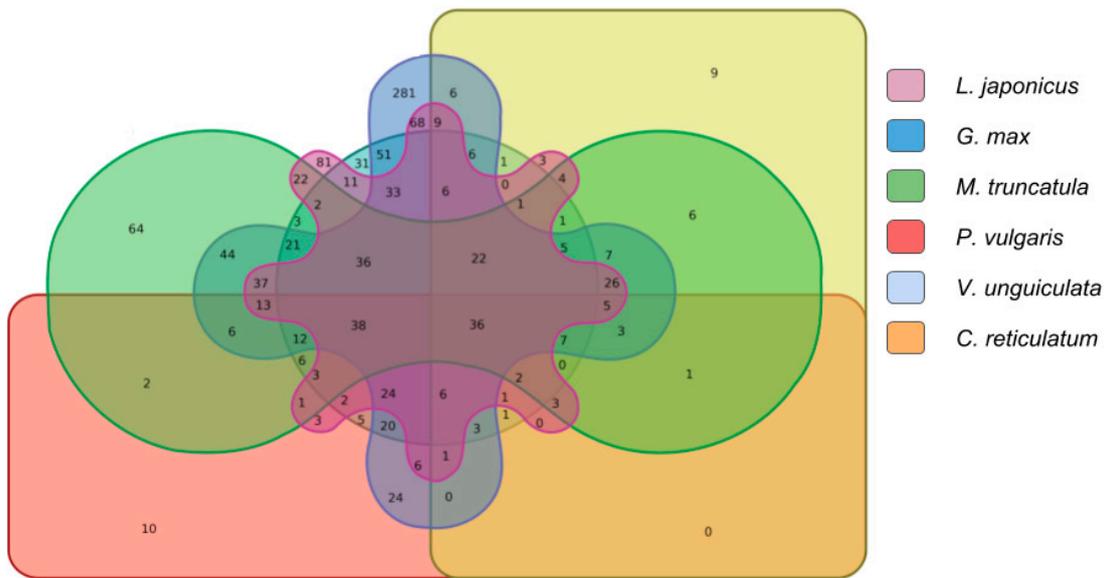


Figure 7. Total transcripts aligned to Fabaceae genomes and transcriptome: This diagram shows the amount of copaiba lncRNA aligned to the both genome and transcriptome of each species segregated by a color pattern indicated in the legend. In the colored overlapped area are the transcripts which were aligned to more than one species genome and transcriptome, and its respective amount.

2.2.4. Expression Analysis of the Conserved Putative lncRNAs among Fabaceae

Most of the lncRNA transcripts that aligned to both genome and transcriptomes of Fabaceae species displayed similar RPKM values for CER and ARF. However, several lncRNAs (256 out of 1141) were regulated at either condition above 2-fold, indicating that they might be involved with the adaptation to different environments (Figure 8). From the total transcripts which aligned to both genome and transcriptome of Fabaceae species, it is possible to identify the formation of two clear groups of putative lncRNAs formed in each condition. The values and identification of the putative lncRNA are available at Table S4.

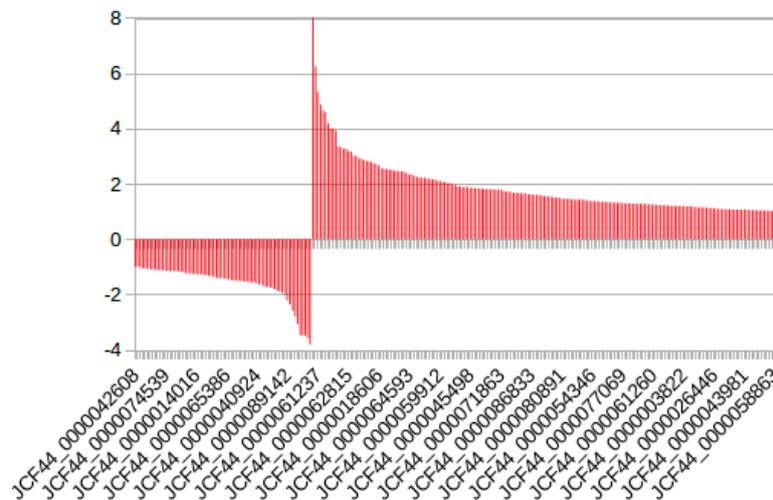


Figure 8. log₂fold change comparison of copaiba lncRNA conserved in Fabaceae species: A total of 1141 lncRNA aligned to multiple Fabaceae genomes and transcriptomes. Comparing the expression of the ARF against CER samples, we found 256 transcripts that presented log₂fc above 1. The graph displays the regulated transcripts of ARF samples compared to CER, each bar corresponds to a single transcript. It is possible to notice that 24 transcripts are regulated above 3-fold.

2.2.5. Identification of Known lncRNA

In order to investigate conserved lncRNAs in closely related species, we used the putative transcripts obtained and compared with BLASTN against CANTATAdb 2.0 and GREnc, which are lncRNA databases for plants. The lncRNAs of *S. bicolor*, *G. max*, *M. truncatula* and *P. vulgaris*, summed up approximately 32,000 sequences. This comparison found a single transcript (JCF45_0000011974/JCF44_0000015840) that matched a *G. max* lncRNA from CANTATA database (CNT2032069) with 90% identity and e-value of 4×10^{-97} . This single transcript is 6.5 fold up-regulated in ARF samples. This transcript was also conserved in *G. max*, *L. japonicus* and *P. vulgaris* genome analysis.

The high sequence identity of the JCF45_0000011974/JCF44_0000015840 transcript with other legume genomes and with a *G. max* lncRNA transcript indicates that it might play biological roles, which could be related to the adaptation to different niches, as it is differentially regulated between CER and ARF populations.

2.3. Stem-Loop Secondary Structure of Regulated Putative lncRNA

Recent studies have suggested that the secondary and even tertiary structures of lncRNAs are conserved and critical for the transcript to be functionally active [77,78]. The secondary structure of lncRNA is regarded as one of the multidimensional conservation pressures that long non-coding transcripts can suffer [79] and as a result of their sequence length, most RNA transcripts are prone to form secondary structure [64]. However, ncRNA present some distinguishable features, such as higher thermostability than coding transcripts, also (ii) their temperature melting (T_m) is significantly higher and (ii) they have greater negative free energy values (minimum free energy—MFE) [79,80]. Another study observed that functional transcripts tend to present higher in silico second structure stability (with greater negative MFE), suggesting a link between secondary structure stability and functionality [81,82].

There are lncRNAs in which the secondary structure dictates their functionality, in humans the lncRNA MEG3 acts as a tumor suppressor based in the structure rather than in the primary sequence conservation [83], also the steroid receptor RNA activator (SRA), one of the few lncRNA that has its secondary structure experimentally defined, is reported to interact with many proteins and be related to breast cancer development. Although SRA primary sequence is mutated, the secondary structure, and the nucleotides involved in the stabilization of the structure are highly conserved, suggesting their direct involvement in the lncRNA functionality [84]. Experimental analysis comparing the folding energy of lncRNAs and mRNAs were capable of differentiating lncRNA based on their higher MFE [79].

Therefore, we performed folding analysis using ViennaRNA (v2.4.8) [85]. In which the stability of secondary structure of RNA can be inferred by MFE, regarding any values below -80 kcal/mol to be structurally stable [77,84]. We selected 256 lncRNAs presenting log2FC above 1 from the set of 1141 conserved lncRNAs in both Fabaceae genome and transcriptome analysis. The majority of lncRNAs (186) presented MFE below -80 (Figure 9). Once these structures were shown to be stable, it is expected that they may have a functional role in which a secondary structure formation is likely to be meaningful [77,79]. At the Table S5, the lncRNA identification, RPKM and predicted MFE values are available.

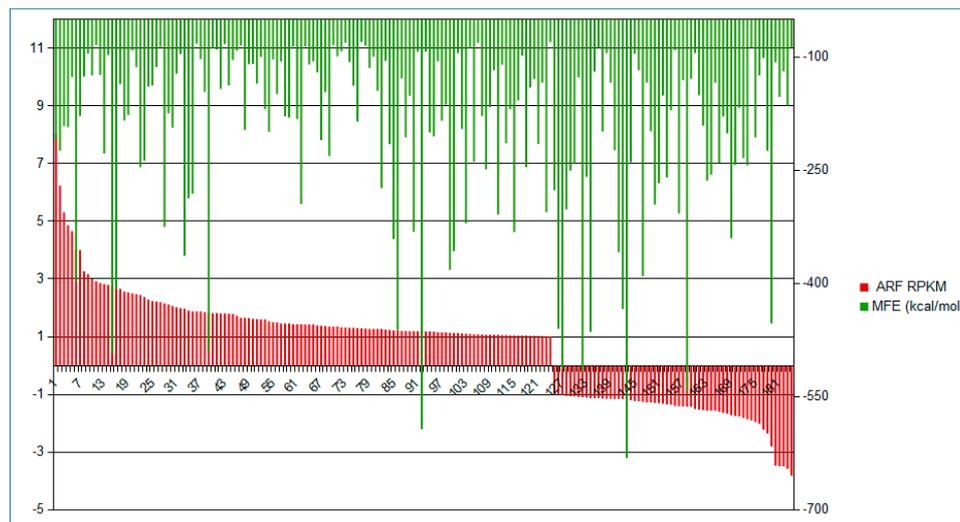


Figure 9. Predicted MFE value and lncRNA relative expression RPKM: The regulated 186 conserved lncRNAs, which were folded to predict their secondary structure stability and presented MFE below -80 kcal/mol. The red bars represent the expression regulation (\log_2FC) of each lncRNA from ARF samples in comparison to CER. The green bars are the transcripts respective MFE value.

2.4. RT-qPCR Analysis of *Copaíba* lncRNA Expression

To validate the expression and sequence of the lncRNA proposed by the RNA-Seq, we selected six conserved lncRNA, considering their high conservation (genome and transcriptome analysis), and regulated expression between the two populations, together with the low MFE value for the predicted structure. The reverse transcription real-time polymerase chain reaction (RT-qPCR) analysis of the lncRNA showed validates the RNA-Seq indication, the transcripts were differentially regulated between the conditions. Since the RNA-Seq was performed from a pool of 10 individuals for each sample, an oscillation was expected regarding the precise differential expression (Figure S1).

2.5. Computational Identification of miRNA and lncRNA Interactions

lncRNA transcripts have been previously reported to interact with microRNA (miRNA), acting as their targets or as a decoy for them [58–62]. In plants, there are tools dedicated to the identification of such interactions, like psRobot [86] and psMimic [61]. Analyzing the 1141 conserved *copaíba* lncRNA transcripts against Fabaceae mature miRNA sequences [87], we identified 94 lncRNA-miRNA interactions with 76 uniquely aligned miRNA. Some of the miRNA families are predicted to bind to more than one lncRNA targets, some of which were regulated in the RNA-Seq, as indicated in Table S6. The lncRNA may also act as endogenous target mimicry (eTM), a novel regulatory pathway in which the non-coding transcript acts as decoy for a miRNA, preventing it from binding to its target, leading to an increased expression of its target mRNA [88–91]. Then, to predict the potential interaction between conserved *copaíba* lncRNA and the known Fabaceae miRNA, we performed a target mimicry analysis through psMimic [61]. In this analysis, 32 lncRNAs transcripts are indicated to potentially act as decoy targets for 31 highly conserved miRNA families which are generally involved in plants' stress response [90,92–94], such as: miR408, miR156, miR164, miR169, miR4392, miR395, miR2673, miR2638, miR7696 and miR11073 (the detailed information is indicated in Table S7).

3. Discussion

Forest trees are a unique group to study adaptability traits, based on their life span and endurance to biotic and abiotic stress [95]. Thus, in the present study we identified 8020 putative lncRNAs, some of which regulated above 2- and 5-fold in either *copaíba* population originated from different biomes. Additionally, a comprehensive analysis of *copaíba* coding transcripts is being conducted by

our group (Franco et al., in preparation), leading to a deeper understanding about their adaptability to these different environments at mRNA and epigenetics level, which is critical to copaíba management and conservation [95,96]. Similarly, Xu et al. [97] performed a study with *Miscanthus lutarioriparius* populations from two different environments and observed an expression profile in which the lncRNA presented higher fold change expression than usually observed with mRNA, suggesting them to be more sensitive and responsive to environmental changes than coding transcripts.

Positional conservation in the genome analysis of lncRNA loci neighboring specific orthologous genes plays an interesting role, particularly when their primary and secondary structures are not completely conserved through species. For example, in humans, a small segment of *AIRN* lncRNA overlaps with the *IGF2R* promoter (insulin-like growth factor 2 receptor) and is sufficient to cause *IGF2R* silencing. This phenomenon was also identified in other lncRNAs [32,34]. Through previous interspecies comparison [29] it was observed that the positional conservation may act by regulating the target gene solely by the transcription of the lncRNA.

In the present copaíba analyses, a total of 22% of the lncRNAs aligned to at least one of the genomes analyzed. The observed amount of lncRNA conservation is coherent with the literature, since the conservation of lncRNA among the same family is much smaller than that of protein-coding genes [25,29,64]. A conservation analysis between human and other placental mammals estimates only 44% of the lncRNA is conserved, while their promoter also seems to be under conservative pressure [98]. In plants, between *Z. mays* and *S. bicolor* there are approximately 25% lncRNA conserved, while *Z. mays* compared to *A. thaliana* presented only 2% conservation [99]. Thus *C. langsdorffii*'s 22% conservation along the Fabaceae family lies in the expected conservation threshold. Within these conserved putative lncRNAs, some were regulated in either condition and it is possible to observe that the majority of them are up-regulated in ARF plants, with a high fold change. The high conservation and regulation among these transcripts suggest that they are under evolutionary constraints, possibly involved in the plant regulatory machinery [32].

Often the lncRNA primary sequence can be degenerated while the position is maintained, indicating that the transcription itself is enough to function as an epigenetic regulator to closely related genes [25,32,64,80,97]. Wu et al. [89] identified an intronic lncRNA, which binds to Curly leaf (CLF) acting as a co-repressor of *AGAMOUS* (*AG*), and it also encodes four new ncRNAs. A study performed comparing putative lncRNA among the *Brassicaceae*, *Aethionemeae* and *Cleomaceae* families uncovered several transcripts that were thought to be lineage-specific, instead they were in fact positionally conserved, although sequence divergent [32]. Moreover, comparing the copaíba lncRNA to several Fabaceae genomes has many advantages, enabling the identification of the transcripts in other species even when they are not being actively transcribed at a given condition. Mapping the lncRNA transcripts to other species cDNA and EST sequences is a way to confirm that it is actively transcribed, and potentially functional [25,26]. Some previous studies that performed this analysis using vertebrates' lncRNA noticed that some transcripts may hold sequence similarities to genomic untranscribed regions [29]. In our study, a higher percentage of lncRNA was paired with EST and cDNA libraries than to the genomic alignment analysis; it might be due to lncRNA processing, which prevents the lncRNA from being correctly mapped to the genome or to the stringent parameters used. The transcripts alignment analysis suggests more orthologous lncRNAs are being actively transcribed throughout the Fabaceae family than initially estimated by genomic analysis.

Regarding that lncRNAs frequently present low expression profile and tendency to be expressed in a tissue or condition specific manner [25,64,80], the transcripts that didn't align to any of the expressed Fabaceae libraries analyzed may still be conserved, although not expressed in the particular condition in which the library samples were taken from. It can be illustrated by the fact that, when comparing copaíba lncRNAs to two databases of putative plants lncRNAs, there was a single transcript that aligned to a known *G. max* lncRNA and mapped to three Fabaceae genomes but didn't align to *G. max* EST and cDNA libraries. Hence, it is essential to understand if the comparison analysis is insufficient to state whether the transcript is expressed or not in a given species.

Moreover, we observed that 14.2% of copaiba lncRNA overlapped the genomic and transcription comparative analysis, while the majority presented similar expression levels in both CER and ARF samples, there were 256 transcripts regulated in either sample above 2-fold. In Figure 8, we can detect a tendency to form two distinct expression profiles among the samples, in which 17 lncRNAs were strongly up-regulated. Considering the samples are from the same type of tissue, and similar developmental stages, it is reasonable to infer that if the lncRNAs identified are functional, the RPKM differential expression between the samples is associated to the plants' response to environmental stimuli. It should be mentioned however that, although there are no replicates in the experiment, RNA samples were prepared from a pool of 10 plants from each local.

During the *in silico* folding prediction of 256 most regulated lncRNAs, which were conserved at genome and transcriptome level, most of the transcripts were regarded to have a stable secondary structure. Thus, a further investigation of these transcripts regarding their functionality is needed, since their structural stability and post-transcriptional processing indicates they are likely to regulate their target expression as molecule, not solely by being transcribed. Thus, the functional characterization and possible targets identification are the natural research directions to further understand the role of lncRNAs in the adaptive response of copaiba and possibly other closely related woody plants.

The miRNA–lncRNA interaction analysis predicted that several of the conserved transcripts might act as potential targets or decoys for miRNA. This interaction has already been observed in other studies [60,61]. The number of putative lncRNA–miRNA targets was similar to Nithin et al. [90] in which the interaction of those ncRNAs was aimed to the crops' improvement. The majority of the miRNA predicted to interact with copaiba transcripts there is reported to be involved in the plants development and response to stress, for instance the miR408 that targets copper protein and plantacyanin genes [100], is responsive to several abiotic stresses [99–105]; the miR156 targets *SPL* transcription factor genes, which are strongly modulated in response to environmental changes [106–109]; the miR164 negatively regulates *NAC* transcription factors during stress [110–114]; also miR169 is extensively studied due to its involvement in the plants response to pathogens infections [94,115,116], and to increase crops resistance to environmental changes [93,117,118].

The identification of the copaiba lncRNAs present an addition to the understanding of lncRNAs in tree plants; several of the transcripts identified are also conserved in other Fabaceae species, which as we understand is the first lncRNA conservational study to analyze this group comparing the genomic and expression data of this plant family. The miRNA–lncRNA predicted interaction presents an interesting cue to the possible role of those lncRNAs in copaiba, however a deeper analysis is necessary to functionally characterize the transcripts.

4. Materials and Methods

4.1. Plant Material Collection

The *Copaifera langsdorffii* leaf samples were collected from two different biomes, ten individual samples were taken from Atlantic Rain Forest (ARF), at Área de Proteção Ambiental da Bacia do Rio São João—Mico Leão, Silva Jardim, RJ, Brazil. And another ten individual samples were collected from Cerrado (CER) ecosystem at Estação Ecológica do Jardim Botânico de Brasília—EEJBB, Distrito Federal, Brazil, in August, during the conspicuous annual dry season. The collected plants specimen was deposited at the Botanical Garden of Rio de Janeiro under the identification number of RB 773246 (CER) and RB 773299 (ARF). All biological material harvested for the expression analysis was placed in RNA later-like buffer, kept at -80°C until extraction.

4.2. RNA Extraction to Sequencing

RNA extraction was performed individually from samples collected from Atlantic Rain Forest and Cerrado. The total RNA was extracted following the modified Japelaghi protocol [119], 10 μg of each RNA samples were sent to Fasteris Life Sciences SA (Plan-les-Ouates, Switzerland), where the quality

and quantity was ascertained by Qubit and Bioanalyzer Nano Chip. Ten individual leaf samples from each ecosystem were selected according to their quality, after passing through poly-A selection protocol. The treated samples were pooled together in equal concentration amounts for sequencing. The ARF and CER samples were named JCF45 and JCF44 respectively, in the multiplex sequencing performed on Illumina HiSeq 2000I using the single-end 76 cycle protocol.

4.3. De Novo Transcriptome Assembly and Gene Expression Profiling

Sequencing reads were checked using FastQC and searched against NCBI nr database using BLASTN (megablast, e-value 1×10^{-5} ; alignment length ≥ 90 and identity $\geq 80\%$) [76]. Reads aligning to non-plant organisms were removed. Raw reads were pre-processed using Trimmomatic (v0.36) [120] for trimming adapters, trailing 15 bases and tailing 5 bases for each read and other quality cleaning. De novo transcriptome assembly was performed using Trinity (v2.3.2; default settings) [121]. Fasta headers were renamed according to the sample ids (e.g., JCF44_xxx and JCF45_xxx). One-to-one correspondence between the transcriptomes was detected by bi-directional BLASTN (v2.2.31+) [76]. Stringent criteria were used for this analysis and only top hits with e-value 1×10^{-10} , query coverage of $\geq 50\%$, identity $\geq 90\%$, bit-score ≥ 50 were considered. Filtered reads were aligned to one-to-one transcripts using Bowtie2 (v. 2.2.9) [68] and RPKM values estimated using Cufflinks (v2.2.1) [69].

4.4. lncRNA Identification

The one-to-one transcripts were screened for lncRNA on several parameters. Transcripts longer than 200 bp with maximum open reading frame (ORF) size of less than 100 amino acids by Getorf (EMBOSS:6.6.0.0) were used to predict putative lncRNAs. These transcripts were analyzed using two software CPC (v.0.9-r2) [66] and PLEK (v1.2) [67] which also filter out transcripts by ORF size and number, transcript length and calculates the transcript coding capacity based on its features. Only transcripts classified as non-coding in both CPC and PLEK analysis were considered as putative lncRNA. Further we kept only those putative lncRNAs with RPKM ≥ 1 .

4.5. Interspecies lncRNAs Conservation Analysis

To identify other lncRNAs which holds sequence conservation among Fabaceae species, we used BLASTN 2.7.1+ (e-value 10^{-15} , identity 90% qcov 50%) [76] against the downloaded GREENC [122] and CANTATAdb v2.0 [63] libraries of Fabaceae lncRNAs, namely *Glycine max*, *Medicago truncatula*, *Phaseolus vulgaris*; both databases present putative lncRNA obtained through their own bioinformatics pipeline, which is similar to the one we used to filter our own.

4.6. Interspecies lncRNA Genome Conservation Analysis

The copaiba putative lncRNA was then mapped through Bowtie2 (v2.3.4.1, default parameters) [68] to seven Fabaceae genomes available at NCBI Genomes (<ftp://ftp.ncbi.nlm.nih.gov/genomes/>): (i) *Vicia faba* (CSVX00000000), (ii) *G. max* (GCF_000004515.4.v2.0) [123], (iii) *M. truncatula* (GCA_000219495.2.v4.0) [71], (iv) *Phaseolus vulgaris* (GCF_000499845.1.v1.0) [72], (v) *Lotus japonica* (GCA_000181115.2.v3.0) [73], (vi) *Vigna unguiculata* (GCA_001687525.1.v0.03) [74] and (vii) *Cicer reticulatum* (GCA_002896235.1.v0.03) [75]. The transcripts which aligned to the aforementioned genomes were compared to assess their conservation in the Fabaceae family using DrawVenn application (<http://bioinformatics.psbuget.be/webtools/Venn/>). It enabled the identification of lncRNAs which mapped to more than one genome.

4.7. Interspecies Conservation of Expressed lncRNA Analysis

The EST and cDNA from Fabaceae species were obtained from PlantGDB [124] and Phytozome [125], namely *C. reticulatum* [75], *P. vulgaris* [72], *M. truncatula* [71], *G. max* [122], *V. unguiculata* [73,74] and *L. japonicus* [73]. The copaiba lncRNA were compared to these Fabaceae

transcripts using BLASTN 2.7.1+ (e-value 10^{-15} , identity 90% qcov 50%) [76] followed by DrawVenn analysis (<http://bioinformatics.psb.ugent.be/webtools/Venn/>) to identify the ones which aligned to more than one species transcriptome. The lncRNA mapped in the genome and transcriptome analysis were selected accordingly to their expression for further analysis.

4.8. Second Structure Modeling

The 1141 copaiba lncRNAs, which were regarded as conserved in the genome and transcriptome analysis, were selected for second structure modeling, using Vienna RNAfold (v2.4.8) [85] package, at 25 °C, default parameters. Subsequently the transcripts which presented second structure stability, evaluated through the MFE value, had their RPKM value compared between the two samples to investigate the potential regulation.

4.9. RT-qPCR Analysis of Copaiba lncRNA Expression

The total RNA was extracted using Japelaghi modified [91] protocol, the total RNA was treated with DNase followed by cDNA syntheses. The copaiba lncRNAs were selected due to their high conservation (transcriptome and genome analysis) and lower MFE value. These six candidates that had their expression analyzed were: JCF44_0000061237 (candidate 1), JCF44_0000046349 (candidate 2), JCF44_0000015840 (candidate 3), JCF44_0000094004 (candidate 4), JCF44_0000040403 (candidate 5) and JCF44_0000021616 (candidate 6). The RT-qPCR was performed using the standard protocol, the primers used are available at Table 2.

Table 2. The following primers were used to assess the relative expression of three selected lncRNAs which presented a higher degree of conservation and predicted second structure stability.

LncRNA ID	Forward Primer Sequence	Reverse Primer Sequence
Candidate 1	AATGCAATACAGCAACCTCTAAACC	GGAGGCACCTGGTGTATTGG
Candidate 2	TCATATCAATGCGGCACTCAA	TGTCTTCAGCTGCCCTTTCTG
Candidate 3	AGCAATTGCGGTTGGTATCC	TGGTACCTTTTCATGTTGCTTTCA
Candidate 4	TCAGGCAGCAGAGGAAGAATC	CACCCAGTTCATGCAACCAA
Candidate 5	CGCCAAATGTCCGCAGAT	GGACTTGCCCGCTATGCA
Candidate 6	AGCAATTGCGGTTGGTATCC	TGGTACCTTTTCATGTTGCTTTCA

4.10. Computational Prediction of miRNA and lncRNA Interaction

The 1141 conserved copaiba lncRNA were analyzed in order to identify its possible role as miRNA targets or their involvement in the miRNA regulatory mechanisms as endogenous targets mimics. The 1960 mature miRNA sequences from nine Fabaceae species were obtained from miRBase (release 22) [87]. The miRNA target prediction was performed locally through psRobot [86] (default parameters, score < 4). The lncRNA–miRNA target mimicry prediction, was performed using psMimic with the default parameters as defined in Wu et al. [61], which considers that: (a) the 2nd to 8th positions at the 5' end of a miRNA must be perfectly aligned to the target, (ii) three unpaired nucleotides are allowed between the 9th to 12th positions at the 5' end of the miRNA sequence, and (iii) at most 3 nucleotide mismatch (excluding bulge region) can be between miRNA and lncRNA sequences.

Supplementary Materials: Supplementary Materials are available online at <http://www.mdpi.com/2311-553X/4/4/27/s1>. Table S1: Presents the fasta sequence of the 8020 putative copaiba lncRNAs transcripts. Table S2: Presents the novel putative copaiba lncRNA identified to be regulated above 2-fold in the comparison of CER and ARF samples. In this table are the identification of the transcripts in each sample, separated by columns and its respective RPKM expression. Table S3: Presents the novel copaiba lncRNA which were expressed in both genome and transcriptome of other Fabaceae species at the conservation analysis. The table presents their identification in each sample, and the respective RPKM expression. Table S4: Presents novel copaiba lncRNA that were expressed in both genome and transcriptome of other Fabaceae species at the conservation analysis, and were regulated above 2-fold between CER and ARF samples, indicating they are potentially involved in the plants adaptive

response to the different biomes. This table presents the identification in each sample, and the respective RPKM expression. Table S5: Presents the identification of a subset of conserved lncRNA, which underwent secondary structural prediction. This file features the RPKM and lncRNAs respective MFE predicted value. Table S6: Table presenting the lncRNA ID and the miRNA which is predicted to target it, provided by psRobot software. Table S7: Table presenting the identification of the lncRNA, which is predicted to act as a decoy target, and the miRNA associated with it. Figure S1: Copaiba lncRNA expression comparison analysis from RNA-Seq and RT-qPCR.

Author Contributions: Conceptualization, P.C.G.F., M.C., A.S.H., S.R.S.C., F.T. and L.O.F.; Methodology, L.O.F., S.R.S.C., M.F.D., T.M.V. and K.C.M.; Software, T.M.V., K.C.M., F.P.; Validation, M.F.D., L.O.F., and S.R.S.C. Formal Analysis, M.F.D.; Investigation, M.F.D.; Resources, S.R.S.C., L.O.F.; Data Curation, M.F.D.; Writing-Original Draft Preparation, M.F.D.; Writing-Review & Editing, M.F.D., F.P., K.C.M., T.M.V.; Visualization, M.F.D.; Supervision, P.C.G.F.; Project Administration, P.C.G.F.; Funding Acquisition, P.C.G.F.

Funding: This research was funded by FAPERJ grant number E12/200.134/2014

Acknowledgments: The authors would like to thank Felipe Casaprima for the support and language revision.

Conflicts of Interest: The authors declare no conflict of interest. The funders had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript, and in the decision to publish the results.

References

1. Junior, V.V.; Pinto, A.C. O GÊNERO *Copaifera* L. *Quim. Nova* **2002**, *25*, 273–286. [[CrossRef](#)]
2. Breitbach, U.B.; Niehues, M.; Lopes, N.P.; Faria, J.E.; Brandão, M.G. Amazonian Brazilian medicinal plants described by CFP von Martius in the 19th century. *J. Ethnopharmacol.* **2013**, *147*, 180–189. [[CrossRef](#)] [[PubMed](#)]
3. Junior, V.V.; Rosas, E.; Carvalho, M.V.D.; Henriques, M.D.G.M.D.O.; Pinto, A.C. Chemical composition and anti-inflammatory activity of copaiba oils from *Copaifera cearensis* Huber ex Ducke, *Copaifera reticulata* Ducke and *Copaifera multijuga* Hayne—A comparative study. *J. Ethnopharmacol.* **2007**, *112*, 248–254.
4. Basile, A.; Sertié, J.; Freitas, P.; Zanini, A. Anti-inflammatory activity of oleoresin from Brazilian *Copaifera*. *J. Ethnopharmacol.* **1988**, *22*, 101–109. [[CrossRef](#)]
5. Falcão, H.D.S.; Lima, I.O.; Santos, V.L.D.; Dantas, H.D.F.; Diniz, M.D.F.; Barbosa-Filho, J.M.; Batista, L.M. Review of the plants with anti-inflammatory activity studied in Brazil. *Rev. Bras. Farmacogn.* **2005**, *15*, 381–391. [[CrossRef](#)]
6. Souza, A.B.; Martins, C.H.; Souza, M.G.; Furtado, N.A.; Heleno, V.C.; de Sousa, J.P.; Rocha, E.M.; Bastos, J.K.; Cunha, W.R.; Veneziani, R.C.; et al. Antimicrobial activity of terpenoids from *Copaifera langsdorffii* Desf. against cariogenic bacteria. *Phytother. Res.* **2011**, *25*, 215–220. [[CrossRef](#)] [[PubMed](#)]
7. Paiva, L.; Rao, V.; Gramosa, N.; Silveira, E. Gastroprotective effect of *Copaifera langsdorffii* oleo-resin on experimental gastric ulcer models in rats. *J. Ethnopharmacol.* **1998**, *62*, 73–78. [[CrossRef](#)]
8. Silva, J.; Borges, V.R.D.A.; Pereira, L.D.C.B.; Ferrari, R.; de Mattos, R.M.; Barros, E.G.D.O.; Palmero, C.Y.; Fernandes, P.D.; de Carvalho, P.R.; Pereira de Sousa, V.; et al. The oil-resin of the tropical rainforest tree *Copaifera langsdorffii* reduces cell viability, changes cell morphology and induces cell death in human endometriotic stromal cultures. *J. Pharm. Pharmacol.* **2015**, *67*, 1744–1755. [[CrossRef](#)] [[PubMed](#)]
9. Masson, D.D.S. *Atividades Cicatrizante e Antimicrobiana do Óleo-Resina de Copaiba (Copaifera langsdorffii) em Úlceras Cutâneas*; Universidade de São Paulo: São Paulo, Brazil, 2011.
10. Gushiken, L.F.S.; Hussni, C.A.; Bastos, J.K.; Rozza, A.L.; Beserra, F.P.; Vieira, A.J.; Padovani, C.R.; Lemos, M.; Polizello Junior, M.; Silva, J.J.M.D.; et al. Skin wound healing potential and mechanisms of the hydroalcoholic extract of leaves and oleoresin of *Copaifera langsdorffii* Desf. Kuntze in rats. *Evid.-Based Complement. Altern. Med.* **2017**, *2017*, 6589270. [[CrossRef](#)] [[PubMed](#)]
11. Millas, A.; Siveira, J.V.; Barbosa, R.; Maria, B.; Koh, I.H.; Bittencourt, E. Implantes de scaffolds de PLGA incorporados com óleo de resina do genero *copaifera* ssp. In Proceedings of the 14^o Congresso da Sociedade Latino Americana de Biomateriais, Orgaos artificiais e Engenharia de Tecidos (SLABO), Maresias, São Paulo, Brazil, 20–24 August 2017.
12. Abed, R.A.; Cavasin, G.M.; Silva, H.H.G.D.; Geris, R.; Silva, I.G.D. Alterações morfohistológicas em larvas de *Aedes aegypti* (Linnaeus, 1762) (Diptera, Culicidae) causadas pela atividade larvicida do óleo-resina da planta medicinal *Copaifera reticulata* Ducke (Leguminosae). *Rev. Patol. Trop.* **2007**, *36*, 75–86. [[CrossRef](#)]

13. Geris, R.; Silva, I.G.D.; Silva, H.H.G.D.; Barison, A.; Rodrigues-Filho, E.; Ferreira, A.G. Diterpenos de *Copaifera reticulata* Ducke com atividade larvívica contra *Aedes aegypti* (L.) (Diptera, Culicidae). *Rev. Inst. Med. Trop. São Paulo* **2008**, *50*, 26–28. [[CrossRef](#)]
14. Santos, A.O.D.; Ueda-Nakamura, T.; Dias Filho, B.P.; Veiga Junior, V.F.; Pinto, A.C.; Nakamura, C.V. Antimicrobial activity of Brazilian copaiba oils obtained from different species of the *Copaifera* genus. *Mem. Inst. Oswaldo Cruz* **2008**, *103*, 277–281. [[CrossRef](#)] [[PubMed](#)]
15. Morelli, C.L.; Mahrous, M.; Belgacem, M.N.; Branciforti, M.C.; Bretas, R.E.S.; Bras, J. Natural copaiba oil as antibacterial agent for bio-based active packaging. *Ind. Crops Prod.* **2015**, *70*, 134–141. [[CrossRef](#)]
16. Guimarães, A.; Cunha, E.A.; Matias, F.O.; Garcia, P.G.; Danopoulos, P.; Swikidisa, R.; Pinheiro, V.A.; Nogueira, R. Antimicrobial Activity of Copaiba (*Copaifera officinalis*) and Pracaxi (*Pentaclethra macroloba*) Oils against *Staphylococcus Aureus*: Importance in Compounding for Wound Care. *Int. J. Pharm. Compd.* **2016**, *20*, 58–62. [[PubMed](#)]
17. Santos, A.O.; Ueda-Nakamura, T.; Dias Filho, B.P.; Junior, V.F.V.; Pinto, A.C.; Nakamura, C.V. Effect of Brazilian copaiba oils on *Leishmania amazonensis*. *J. Ethnopharmacol.* **2008**, *120*, 204–208. [[CrossRef](#)] [[PubMed](#)]
18. Calderon, L.D.A.; Silva-Jardim, I.; Zuliani, J.P.; Ciancaglini, P.; Silva, L.H.P.D.; Stábeli, R.G. Amazonian biodiversity: A view of drug development for leishmaniasis and malaria. *J. Braz. Chem. Soc.* **2009**, *20*, 1011–1023. [[CrossRef](#)]
19. Soares, D.C.; Portella, N.A.; Ramos, M.F.D.S.; Siani, A.C.; Saraiva, E.M. Trans- β -caryophyllene: An effective antileishmanial compound found in commercial copaiba oil (*Copaifera* spp.). *Evid.-Based Complement. Altern. Med.* **2013**, *2013*, 761323. [[CrossRef](#)] [[PubMed](#)]
20. Mizuno, C.S.; Souza, A.B.; Tekwani, B.L.; Ambrósio, S.R.; Veneziani, R.C. Synthesis and biological evaluation of polyalthic acid derivatives for the treatment of neglected diseases. *Bioorg. Med. Chem. Lett.* **2015**, *25*, 5529–5531. [[CrossRef](#)] [[PubMed](#)]
21. de Moraes, A.R.D.P.; Tavares, G.D.; Rocha, F.J.S.; de Paula, E.; Giorgio, S. Effects of nanoemulsions prepared with essential oils of copaiba-and andiroba against *Leishmania infantum* and *Leishmania amazonensis* infections. *Exp. Parasitol.* **2018**, *187*, 12–21. [[CrossRef](#)] [[PubMed](#)]
22. Sâmia, R.; de Oliveira, R.; Moscardini, V.; Carvalho, G. Effects of aqueous extracts of *Copaifera langsdorffii* (Fabaceae) on the growth and reproduction of *Spodoptera frugiperda* (JE Smith) (Lepidoptera: Noctuidae). *Neotrop. Entomol.* **2016**, *45*, 580–587. [[CrossRef](#)] [[PubMed](#)]
23. Alencar, J.D.C. Estudos silviculturais de uma população natural de *Copaifera multijuga* Hayne-Leguminosae, na Amazônia Central. 2-Produção de óleo-resina. *Acta Amazon.* **1982**, *12*, 75–89. [[CrossRef](#)]
24. Carvalho, M.C.; Gomide, L.R.; Santos, R.M.D.; Scolforo, J.R.S.; Carvalho, L.M.T.D.; Mello, J.M.D. Modeling ecological niche of tree species in Brazilian tropical area. *Cerne* **2017**, *23*, 229–240. [[CrossRef](#)]
25. Johnsson, P.; Lipovich, L.; Grandér, D.; Morris, K.V. Evolutionary conservation of long non-coding RNAs; sequence, structure, function. *Biochim. Biophys. Acta (BBA)-Gen. Subj.* **2014**, *1840*, 1063–1071. [[CrossRef](#)] [[PubMed](#)]
26. Kapusta, A.; Feschotte, C. Volatile evolution of long noncoding RNA repertoires: Mechanisms and biological implications. *Trends Genet.* **2014**, *30*, 439–452. [[CrossRef](#)] [[PubMed](#)]
27. Ma, X.; Shao, C.; Jin, Y.; Wang, H.; Meng, Y. Long non-coding RNAs: A novel endogenous source for the generation of Dicer-like 1-dependent small RNAs in *Arabidopsis thaliana*. *RNA Biol.* **2014**, *11*, 373–390. [[CrossRef](#)] [[PubMed](#)]
28. Ariel, F.; Romero-Barrios, N.; Jégu, T.; Benhamed, M.; Crespi, M. Battles and hijacks: Noncoding transcription in plants. *Trends Plant Sci.* **2015**, *20*, 362–371. [[CrossRef](#)] [[PubMed](#)]
29. Hezroni, H.; Koppstein, D.; Schwartz, M.G.; Avrutin, A.; Bartel, D.P.; Ulitsky, I. Principles of long noncoding RNA evolution derived from direct comparison of transcriptomes in 17 species. *Cell Rep.* **2015**, *11*, 1110–1122. [[CrossRef](#)] [[PubMed](#)]
30. Kashi, K.; Henderson, L.; Bonetti, A.; Carninci, P. Discovery and functional analysis of lncRNAs: Methodologies to investigate an uncharacterized transcriptome. *Biochim. Biophys. Acta (BBA)-Gene Regul. Mech.* **2016**, *1859*, 3–15. [[CrossRef](#)] [[PubMed](#)]
31. Bhatia, G.; Goyal, N.; Sharma, S.; Upadhyay, S.K.; Singh, K. Present scenario of long non-coding RNAs in plants. *Non-Coding RNA* **2017**, *3*, 16. [[CrossRef](#)] [[PubMed](#)]

32. Mohammadin, S.; Edger, P.P.; Pires, J.C.; Schranz, M.E. Positionally-conserved but sequence-diverged: Identification of long non-coding RNAs in the Brassicaceae and Cleomaceae. *BMC Plant Biol.* **2015**, *15*, 217. [[CrossRef](#)] [[PubMed](#)]
33. Shafiq, S.; Li, J.; Sun, Q. Functions of plants long non-coding RNAs. *Biochim. Biophys. Acta (BBA)-Gene Regul. Mech.* **2016**, *1859*, 155–162. [[CrossRef](#)] [[PubMed](#)]
34. Guttman, M.; Donaghey, J.; Carey, B.W.; Garber, M.; Grenier, J.K.; Munson, G.; Young, G.; Lucas, A.B.; Ach, R.; Bruhn, L.; et al. lincRNAs act in the circuitry controlling pluripotency and differentiation. *Nature* **2011**, *477*, 295. [[CrossRef](#)] [[PubMed](#)]
35. Kitagawa, M.; Kotake, Y.; Ohhata, T. Long non-coding RNAs involved in cancer development and cell fate determination. *Curr. Drug Targets* **2012**, *13*, 1616–1621. [[CrossRef](#)] [[PubMed](#)]
36. Bardou, F.; Ariel, F.; Simpson, C.G.; Romero-Barrios, N.; Laporte, P.; Balzergue, S.; Brown, J.W.; Crespi, M. Long noncoding RNA modulates alternative splicing regulators in *Arabidopsis*. *Dev. Cell* **2014**, *30*, 166–176. [[CrossRef](#)] [[PubMed](#)]
37. Bazin, J.; Bailey-Serres, J. Emerging roles of long non-coding RNA in root developmental plasticity and regulation of phosphate homeostasis. *Front. Plant Sci.* **2015**, *6*, 400. [[CrossRef](#)] [[PubMed](#)]
38. Liu, X.; Hao, L.; Li, D.; Zhu, L.; Hu, S. Long non-coding RNAs and their biological roles in plants. *Genom. Proteom. Bioinform.* **2015**, *13*, 137–147. [[CrossRef](#)] [[PubMed](#)]
39. Amor, B.B.; Wirth, S.; Merchan, F.; Laporte, P.; d'Aubenton Carafa, Y.; Hirsch, J.; Maizel, A.; Mallory, A.; Lucas, A.; Deragon, J.M.; et al. Novel long non-protein coding RNAs involved in *Arabidopsis* differentiation and stress responses. *Genome Res.* **2008**, *19*, 57–69. [[CrossRef](#)] [[PubMed](#)]
40. Chen, M.; Wang, C.; Bao, H.; Chen, H.; Wang, Y. Genome-wide identification and characterization of novel lncRNAs in *Populus* under nitrogen deficiency. *Mol. Genet. Genom.* **2016**, *291*, 1663–1680. [[CrossRef](#)] [[PubMed](#)]
41. Gai, Y.-P.; Yuan, S.-S.; Zhao, Y.-N.; Zhao, H.-N.; Zhang, H.-L.; Ji, X.-L. A Novel LncRNA, MuLnc1, Associated with Environmental Stress in Mulberry (*Morus multicaulis*). *Front. Plant Sci.* **2018**, *9*, 669. [[CrossRef](#)] [[PubMed](#)]
42. He, Y. Chromatin regulation of flowering. *Trends Plant Sci.* **2012**, *17*, 556–562. [[CrossRef](#)] [[PubMed](#)]
43. Zhu, Y.; Rowley, M.J.; Böhmendorfer, G.; Wierzbicki, A.T. A SWI/SNF chromatin-remodeling complex acts in noncoding RNA-mediated transcriptional silencing. *Mol. Cell* **2013**, *49*, 298–309. [[CrossRef](#)] [[PubMed](#)]
44. Vance, K.W.; Ponting, C.P. Transcriptional regulatory functions of nuclear long noncoding RNAs. *Trends Genet.* **2014**, *30*, 348–355. [[CrossRef](#)] [[PubMed](#)]
45. Böhmendorfer, G.; Wierzbicki, A.T. Control of chromatin structure by long noncoding RNA. *Trends Cell Biol.* **2015**, *25*, 623–632. [[CrossRef](#)] [[PubMed](#)]
46. Ariel, F.; Jegu, T.; Latrasse, D.; Romero-Barrios, N.; Christ, A.; Benhamed, M.; Crespi, M. Noncoding transcription by alternative RNA polymerases dynamically regulates an auxin-driven chromatin loop. *Mol. Cell* **2014**, *55*, 383–396. [[CrossRef](#)] [[PubMed](#)]
47. Wang, K.C.; Chang, H.Y. Molecular mechanisms of long noncoding RNAs. *Mol. Cell* **2011**, *43*, 904–914. [[CrossRef](#)] [[PubMed](#)]
48. Kim, E.-D.; Sung, S. Long noncoding RNA: Unveiling hidden layer of gene regulatory networks. *Trends Plant Sci.* **2012**, *17*, 16–21. [[CrossRef](#)] [[PubMed](#)]
49. Heo, J.B.; Sung, S. Vernalization-mediated epigenetic silencing by a long intronic noncoding RNA. *Science* **2011**, *331*, 76–79. [[CrossRef](#)] [[PubMed](#)]
50. Csorba, T.; Questa, J.I.; Sun, Q.; Dean, C. Antisense COOLAIR mediates the coordinated switching of chromatin states at FLC during vernalization. *Proc. Natl. Acad. Sci. USA* **2014**, *111*, 16160–16165. [[CrossRef](#)] [[PubMed](#)]
51. Kim, D.-H.; Sung, S. Vernalization-triggered intragenic chromatin loop formation by long noncoding RNAs. *Dev. Cell* **2017**, *40*, 302–312. [[CrossRef](#)] [[PubMed](#)]
52. Ding, J.; Lu, Q.; Ouyang, Y.; Mao, H.; Zhang, P.; Yao, J.; Xu, C.; Li, X.; Xiao, J.; Zhang, Q. A long noncoding RNA regulates photoperiod-sensitive male sterility, an essential component of hybrid rice. *Proc. Natl. Acad. Sci. USA* **2012**, *109*, 2654–2659. [[CrossRef](#)] [[PubMed](#)]
53. Jiao, Y.; Lau, O.S.; Deng, X.W. Light-regulated transcriptional networks in higher plants. *Nat. Rev. Genet.* **2007**, *8*, 217. [[CrossRef](#)] [[PubMed](#)]
54. Shi, W.; Quan, M.; Du, Q.; Zhang, D. The interactions between the long non-coding RNA NERDL and its target gene affect wood formation in *Populus tomentosa*. *Front. Plant Sci.* **2017**, *8*, 1035. [[CrossRef](#)] [[PubMed](#)]

55. Jabnourne, M.; Secco, D.; Lecampion, C.; Robaglia, C.; Shu, Q.; Poirier, Y. A rice cis-natural antisense RNA acts as a translational enhancer for its cognate mRNA and contributes to phosphate homeostasis and plant fitness. *Plant Cell* **2013**, tpc-113. [[CrossRef](#)] [[PubMed](#)]
56. Zhang, W.; Han, Z.; Guo, Q.; Liu, Y.; Zheng, Y.; Wu, F.; Jin, W. Identification of maize long non-coding RNAs responsive to drought stress. *PLoS ONE* **2014**, *9*, e98958. [[CrossRef](#)] [[PubMed](#)]
57. Lv, Y.; Liang, Z.; Ge, M.; Qi, W.; Zhang, T.; Lin, F.; Peng, Z.; Zhao, H. Genome-wide identification and functional prediction of nitrogen-responsive intergenic and intronic long non-coding RNAs in maize (*Zea mays* L.). *BMC Genom.* **2016**, *17*, 350. [[CrossRef](#)] [[PubMed](#)]
58. Sunkar, R.; Zhu, J.-K. Novel and stress-regulated microRNAs and other small RNAs from *Arabidopsis*. *Plant Cell* **2004**, *16*, 2001–2019. [[CrossRef](#)] [[PubMed](#)]
59. Arikiti, S.; Zhai, J.; Meyers, B.C. Biogenesis and function of rice small RNAs from non-coding RNA precursors. *Curr. Opin. Plant Biol.* **2013**, *16*, 170–179. [[CrossRef](#)] [[PubMed](#)]
60. Franco-Zorrilla, J.M.; Valli, A.; Todesco, M.; Mateos, I.; Puga, M.I.; Rubio-Somoza, I.; Leyva, A.; Weigel, D.; Garcia, J.A.; Paz-Ares, J. Target mimicry provides a new mechanism for regulation of microRNA activity. *Nat. Genet.* **2007**, *39*, 1033. [[CrossRef](#)] [[PubMed](#)]
61. Wu, H.-J.; Wang, Z.-M.; Wang, M.; Wang, X.-J. Widespread long noncoding RNAs as endogenous target mimics for microRNAs in plants. *Plant Physiol.* **2013**, *161*, 1875–1884. [[CrossRef](#)] [[PubMed](#)]
62. Wang, J.; Yu, W.; Yang, Y.; Li, X.; Chen, T.; Liu, T.; Ma, N.; Yang, X.; Liu, R.; Zhang, B. Genome-wide analysis of tomato long non-coding RNAs and identification as endogenous target mimic for microRNA in response to TYLCV infection. *Sci. Rep.* **2015**, *5*, 16946. [[CrossRef](#)] [[PubMed](#)]
63. Szcześniak, M.W.; Rosikiewicz, W.; Makalowska, I. CANTATAdb: A collection of plant long non-coding RNAs. *Plant Cell Physiol.* **2015**, *57*, e8. [[CrossRef](#)] [[PubMed](#)]
64. Ulitsky, I. Evolution to the rescue: Using comparative genomics to understand long non-coding RNAs. *Nat. Rev. Genet.* **2016**, *17*, 601. [[CrossRef](#)] [[PubMed](#)]
65. Hezroni, H.; Perry, R.B.-T.; Meir, Z.; Housman, G.; Lubelsky, Y.; Ulitsky, I. A subset of conserved mammalian long non-coding RNAs are fossils of ancestral protein-coding genes. *Genome Biol.* **2017**, *18*, 162. [[CrossRef](#)] [[PubMed](#)]
66. Kong, L.; Zhang, Y.; Ye, Z.-Q.; Liu, X.-Q.; Zhao, S.-Q.; Wei, L.; Gao, G. CPC: Assess the protein-coding potential of transcripts using sequence features and support vector machine. *Nucleic Acids Res.* **2007**, *35*, W345–W349. [[CrossRef](#)] [[PubMed](#)]
67. Li, A.; Zhang, J.; Zhou, Z. PLEK: A tool for predicting long non-coding RNAs and messenger RNAs based on an improved k-mer scheme. *BMC Bioinform.* **2014**, *15*, 311. [[CrossRef](#)] [[PubMed](#)]
68. Langmead, B.; Salzberg, S.L. Fast gapped-read alignment with Bowtie 2. *Nat. Methods* **2012**, *9*, 357. [[CrossRef](#)] [[PubMed](#)]
69. Trapnell, C.; Roberts, A.; Goff, L.; Pertea, G.; Kim, D.; Kelley, D.R.; Pimentel, H.; Salzberg, S.L.; Rinn, J.L.; Pachter, L. Differential gene and transcript expression analysis of RNA-seq experiments with TopHat and Cufflinks. *Nat. Protocols* **2012**, *7*, 562. [[CrossRef](#)] [[PubMed](#)]
70. Schmutz, J.; Cannon, S.B.; Schlueter, J.; Ma, J.; Mitros, T.; Nelson, W.; Hyten, D.L.; Song, Q.; Thelen, J.J.; Cheng, J.; et al. Genome sequence of the palaeopolyploid soybean. *Nature* **2010**, *463*, 178. [[CrossRef](#)] [[PubMed](#)]
71. Tang, H.; Krishnakumar, V.; Bidwell, S.; Rosen, B.; Chan, A.; Zhou, S.; Gentzmittel, L.; Childs, K.L.; Yandell, M.; Gundlach, H.; et al. An improved genome release (version Mt4. 0) for the model legume *Medicago truncatula*. *BMC Genom.* **2014**, *15*, 312. [[CrossRef](#)] [[PubMed](#)]
72. Vlasova, A.; Capella-Gutiérrez, S.; Rendón-Anaya, M.; Hernández-Oñate, M.; Minoche, A.E.; Erb, I.; Câmara, F.; Prieto-Barja, P.; Corvelo, A.; Sanseverino, W.; et al. Genome and transcriptome analysis of the Mesoamerican common bean and the role of gene duplications in establishing tissue and temporal specialization of genes. *Genome Biol.* **2016**, *17*, 32. [[CrossRef](#)] [[PubMed](#)]
73. Sato, S.; Nakamura, Y.; Kaneko, T.; Asamizu, E.; Kato, T.; Nakao, M.; Sasamoto, S.; Watanabe, A.; Ono, A.; Kawashima, K.; et al. Genome structure of the legume, *Lotus japonicus*. *DNA Res.* **2008**, *15*, 227–239. [[CrossRef](#)] [[PubMed](#)]
74. Munoz-Amatriain, M.; Mirebrahim, H.; Xu, P.; Wanamaker, S.I.; Luo, M.; Alhakami, H.; Alpert, M.; Atokple, I.; Batieno, B.J.; Boukar, O.; et al. Genome resources for climate-resilient cowpea, an essential crop for food security. *Plant J.* **2017**, *89*, 1042–1054. [[CrossRef](#)] [[PubMed](#)]

75. Gupta, S.; Nawaz, K.; Parween, S.; Roy, R.; Sahu, K.; Kumar Pole, A.; Khandal, H.; Srivastava, R.; Kumar Parida, S.; Chattopadhyay, D. Draft genome sequence of *Cicer reticulatum* L., the wild progenitor of chickpea provides a resource for agronomic trait improvement. *DNA Res.* **2016**, *24*, 1–10.
76. Camacho, C.; Coulouris, G.; Avagyan, V.; Ma, N.; Papadopoulos, J.; Bealer, K.; Madden, T.L. BLAST+: Architecture and applications. *BMC Bioinform.* **2009**, *10*, 421. [[CrossRef](#)] [[PubMed](#)]
77. Novikova, I.V.; Hennelly, S.P.; Sanbonmatsu, K.Y. Sizing up long non-coding RNAs: Do lncRNAs have secondary and tertiary structure? *BioArchitecture* **2012**, *2*, 189–199. [[CrossRef](#)] [[PubMed](#)]
78. Diederichs, S. The four dimensions of noncoding RNA conservation. *Trends Genet.* **2014**, *30*, 121–123. [[CrossRef](#)] [[PubMed](#)]
79. Kertesz, M.; Wan, Y.; Mazor, E.; Rinn, J.L.; Nutter, R.C.; Chang, H.Y.; Segal, E. Genome-wide measurement of RNA secondary structure in yeast. *Nature* **2010**, *467*, 103–107. [[CrossRef](#)] [[PubMed](#)]
80. Mercer, R.; Mattick, J.S. Structure and Function of long noncoding RNAs in epigenetic regulation. *Nat. Struct. Mol. Biol.* **2013**, *20*, 300–307. [[CrossRef](#)] [[PubMed](#)]
81. Steffens, S.; Thiel, H.-J.; Behrens, S.-E. The RNA-Dependent RNA polymerases of different members of the family flaviviridae exhibit similar properties in vitro. *J. Gen. Virol.* **1999**, *80*, 2583–2590. [[CrossRef](#)] [[PubMed](#)]
82. Wan, Y.; Qu, K.; Ouyang, Z.; Kertesz, M.; Li, J.; Tibshirani, R.; Makino, D.; Nutter, R.; Segal, E.; Chang, H. Genome-wide measurement of RNA folding energies. *Mol. Cell* **2012**, *48*, 169–181. [[CrossRef](#)] [[PubMed](#)]
83. Zhang, X.; Rice, K.; Wang, Y.; Chen, W.; Zhong, Y.; Nakayama, Y.; Zhou, Y.; Klibanski, A. Maternally expressed gene 3 (*MEG3*) noncoding ribonucleic acid: Isoform structure, expression, and functions. *Endocrinology* **2010**, *151*, 939–947. [[CrossRef](#)] [[PubMed](#)]
84. Novikova, I.; Hennelly, S.P.; Sanbonmatsu, K. Structural architecture of the human long non-coding RNA, steroid receptor RNA activator. *Nucleic Acids Res.* **2012**, *40*, 5034–5051. [[CrossRef](#)] [[PubMed](#)]
85. Lorenz, R.; Bernhart, S.; Höner zu Siederdissen, C.; Tafer, H.; Flamm, C.; Stadler, P.; Hofacker, I. ViennaRNA Package 2.0. *Algorithms Mol. Biol.* **2011**, *6*, 26. [[CrossRef](#)] [[PubMed](#)]
86. Wu, H.-J.; Ma, Y.-K.; Chen, T.; Wang, M.; Wang, X.-J. PsRobot: A web-based plant small RNA meta-analysis toolbox. *Nucleic Acids Res.* **2012**, *40*, W22–W28. [[CrossRef](#)] [[PubMed](#)]
87. Griffiths-Jones, S.; Grocock, R.J.; Van Dongen, S.; Bateman, A.; Enright, A.J. miRBase: MicroRNA sequences, targets and gene nomenclature. *Nucleic Acids Res.* **2006**, *34*, D140–D144. [[CrossRef](#)] [[PubMed](#)]
88. Karakulah, G.; Kurtoğlu, K.Y.; Unver, T. PeTmBase: A database of plant endogenous target mimics (eTMs). *PLoS ONE* **2016**, *11*, e0167698. [[CrossRef](#)] [[PubMed](#)]
89. Wu, H.W.; Deng, S.; Xu, H.; Mao, H.Z.; Liu, J.; Niu, Q.W.; Wang, H.; Chua, N.H. A noncoding RNA transcribed from the AGAMOUS (AG) second intron binds to CURLY LEAF and represses AG expression in leaves. *New Phytol.* **2018**. [[CrossRef](#)] [[PubMed](#)]
90. Nithin, C.; Thomas, A.; Basak, J.; Bahadur, R.P. Genome-wide identification of miRNAs and lncRNAs in *Cajanus cajan*. *BMC Genom.* **2017**, *18*, 878. [[CrossRef](#)] [[PubMed](#)]
91. Axtell, M.J. Lost in translation? microRNAs at the rough ER. *Trends Plant Sci.* **2017**, *22*, 273–274. [[CrossRef](#)] [[PubMed](#)]
92. Yuan, C.; Meng, X.; Li, X.; Illing, N.; Ingle, R.A.; Wang, J.; Chen, M. PceRBase: A database of plant competing endogenous RNA. *Nucleic Acids Res.* **2017**, *45*, D1009–D1014. [[CrossRef](#)] [[PubMed](#)]
93. Zhang, B. MicroRNA: A new target for improving plant tolerance to abiotic stress. *J. Exp. Bot.* **2015**, *66*, 1749–1761. [[CrossRef](#)] [[PubMed](#)]
94. Hoang, T.V.; Vo, K.T.X.; Hong, W.-J.; Jung, K.-H.; Jeon, J.-S. Defense response to pathogens through epigenetic regulation in rice. *J. Plant Biol.* **2018**, *61*, 1–10. [[CrossRef](#)]
95. Wang, L.; Xia, X.; Jiang, H.; Lu, Z.; Cui, J.; Cao, F.; Jin, B. Genome-wide identification and characterization of novel NAs in *Ginkgo biloba*. *Trees* **2018**, *32*, 1429–1442. [[CrossRef](#)]
96. Aitken, S.N.; Yeaman, S.; Holliday, J.A.; Wang, T.; Curtis-McLane, S. Adaptation, migration or extirpation: Climate change outcomes for tree populations. *Evol. Appl.* **2008**, *1*, 95–111. [[CrossRef](#)] [[PubMed](#)]
97. Xu, Q.; Song, Z.; Zhu, C.; Tao, C.; Kang, L.; Liu, W.; He, F.; Yan, J.; Sang, T. Systematic comparison of lncRNAs with protein coding MRNAs in population expression and their response to environmental change. *BMC Plant Biol.* **2017**, *17*, 42. [[CrossRef](#)] [[PubMed](#)]
98. Derrien, T.; Johnson, R.; Bussotti, G.; Tanzer, A.; Djebali, S.; Tilgner, H.; Guernec, G.; Martin, D.; Merkel, A.; Knowles, D.G.; et al. The GENCODE v7 catalog of human long noncoding RNAs: Analysis of their gene structure, evolution, and expression. *Genome Res.* **2012**, *22*, 1775–1789. [[CrossRef](#)] [[PubMed](#)]

99. Li, L.; Eichten, S.R.; Shimizu, R.; Petsch, K.; Yeh, C.T.; Wu, W.; Chetoor, A.M.; Givan, S.A.; Cole, R.A.; Fowler, J.E.; et al. Genome-wide discovery and characterization of maize long non-coding RNAs. *Genome Biol.* **2014**, *15*, 40. [[CrossRef](#)] [[PubMed](#)]
100. Trindade, I.; Capitão, C.; Dalmay, T.; Fevereiro, M.P.; Dos Santos, D.M. miR398 and miR408 are up-regulated in response to water deficit in *Medicago truncatula*. *Planta* **2010**, *231*, 705–716. [[CrossRef](#)] [[PubMed](#)]
101. Feng, H.; Zhang, Q.; Wang, Q.; Wang, X.; Liu, J.; Li, M.; Huang, L.; Kang, Z. Target of tae-miR408, a chemocyanin-like protein gene (*TaCLP1*), plays positive roles in wheat response to high-salinity, heavy cupric stress and stripe rust. *Plant Mol. Biol.* **2013**, *83*, 433–443. [[CrossRef](#)] [[PubMed](#)]
102. Hajyzadeh, M.; Turktas, M.; Khawar, K.M.; Unver, T. miR408 overexpression causes increased drought tolerance in chickpea. *Gene* **2015**, *555*, 186–193. [[CrossRef](#)] [[PubMed](#)]
103. Zhao, X.Y.; Hong, P.; Wu, J.Y.; Chen, X.B.; Ye, X.G.; Pan, Y.Y.; Wang, J.; Zhang, X.S. The tae-miR408-mediated control of *TaTOC1* gene transcription is required for the regulation of heading time in wheat (*Triticum aestivum* L.). *Plant Physiol.* **2016**. [[CrossRef](#)] [[PubMed](#)]
104. Zhang, J.-P.; Yu, Y.; Feng, Y.-Z.; Zhou, Y.-F.; Zhang, F.; Yang, Y.-W.; Lei, M.-Q.; Zhang, Y.-C.; Chen, Y.-Q. miRNA MiR408 regulates grain yield and photosynthesis via a phytoecyanin protein. *Plant Physiol.* **2017**. [[CrossRef](#)] [[PubMed](#)]
105. Song, Z.; Zhang, L.; Wang, Y.; Li, H.; Li, S.; Zhao, H.; Zhang, H. Constitutive expression of miR408 improves biomass and seed yield in *Arabidopsis*. *Front. Plant Sci.* **2018**, *8*, 2114. [[CrossRef](#)] [[PubMed](#)]
106. Stief, A.; Altmann, S.; Hoffmann, K.; Pant, B.D.; Scheible, W.-R.; Bäurle, I. *Arabidopsis* miR156 regulates tolerance to recurring environmental stress through SPL transcription factors. *Plant Cell* **2014**, tpc-114. [[CrossRef](#)] [[PubMed](#)]
107. Yu, N.; Niu, Q.-W.; Ng, K.-H.; Chua, N.-H. The role of miR156/SPL s modules in *Arabidopsis* lateral root development. *Plant J.* **2015**, *83*, 673–685. [[CrossRef](#)] [[PubMed](#)]
108. Guo, C.; Xu, Y.; Shi, M.; Lai, Y.; Wu, X.; Wang, H.; Zhu, Z.; Poethig, R.S.; Wu, G. Repression of miR156 by miR159 regulates the timing of the juvenile-to-adult transition in *Arabidopsis*. *Plant Cell* **2017**, tpc-00975. [[CrossRef](#)] [[PubMed](#)]
109. Ge, Y.; Han, J.; Zhou, G.; Xu, Y.; Ding, Y.; Shi, M.; Guo, C.; Wu, G. Silencing of miR156 confers enhanced resistance to brown planthopper in rice. *Planta* **2018**, *248*, 813–826. [[CrossRef](#)] [[PubMed](#)]
110. Fang, Y.; Xie, K.; Xiong, L. Conserved miR164-targeted NAC genes negatively regulate drought resistance in rice. *J. Exp. Bot.* **2014**, *65*, 2119–2135. [[CrossRef](#)] [[PubMed](#)]
111. Xiong, L.; Xie, K.; Hou, X. MIR164 Gene That Controls Plant Root System Development and Fertility and Use Thereof. U.S. Patent No. 8,802,929, 12 August 2014.
112. Jun, W.; Junhong, Z.; Menghui, H.; Minhui, Z.; Zaikang, T. Expression analysis of miR164 and its target gene *NAC1* in response to low nitrate availability in *Betula luminifera*. *Hereditas* **2015**, *38*, 155–162.
113. Li, J.; Lai, T.; Song, H.; Xu, X. miR164 is involved in delaying senescence of strawberry (*Fragaria ananassa*) fruit by negatively regulating NAC transcription factor genes under low temperature. *Russ. J. Plant Physiol.* **2017**, *64*, 251–259. [[CrossRef](#)]
114. Lu, X.; Dun, H.; Lian, C.; Zhang, X.; Yin, W.; Xia, X. The role of peu-miR164 and its target *PeNAC* genes in response to abiotic stress in *Populus euphratica*. *Plant Physiol. Biochem.* **2017**, *115*, 418–438. [[CrossRef](#)] [[PubMed](#)]
115. Zhao, C.; Xia, H.; Cao, T.; Yang, Y.; Zhao, S.; Hou, L.; Zhang, Y.; Li, C.; Zhang, X.; Wang, X. Small RNA and degradome deep sequencing reveals peanut microRNA roles in response to pathogen infection. *Plant Mol. Biol. Rep.* **2015**, *33*, 1013–1029. [[CrossRef](#)]
116. Hanemian, M.; Barlet, X.; Sorin, C.; Yadeta, K.A.; Keller, H.; Favery, B.; Simon, R.; Thomma, B.P.; Hartmann, C.; Crespi, M.; et al. *Arabidopsis* CLAVATA1 and CLAVATA2 receptors contribute to *Ralstonia solanacearum* pathogenicity through a miR169-dependent pathway. *New Phytol.* **2016**, *211*, 502–515. [[CrossRef](#)] [[PubMed](#)]
117. Serivichyaswat, P.T.; Susila, H.; Ahn, J.H. Elongated hypocotyl 5-homolog (HYH) negatively regulates expression of the ambient temperature-responsive microRNA gene *MIR169*. *Front. Plant Sci.* **2017**, *8*, 2087. [[CrossRef](#)] [[PubMed](#)]
118. Moradi, K.; Khalili, F. Assessment of pattern expression of miR172 and miR169 in response to drought stress in *Echinacea purpurea* L. *Biocatal. Agric. Biotechnol.* **2018**, in press. [[CrossRef](#)]

119. Japelaghi, R.H.; Haddad, R.; Garoosi, G.A. Rapid and efficient isolation of high quality nucleic acids from plant tissues rich in polyphenols and polysaccharides. *Mol. Biotechnol.* **2011**, *49*, 129–137. [[CrossRef](#)] [[PubMed](#)]
120. Bolger, A.M.; Lohse, M.; Usadel, B. Trimmomatic: A flexible trimmer for Illumina sequence data. *Bioinformatics* **2014**, *30*, 2114–2120. [[CrossRef](#)] [[PubMed](#)]
121. Grabherr, M.G.; Haas, B.J.; Yassour, M.; Levin, J.Z.; Thompson, D.A.; Amit, I.; Adiconis, X.; Fan, L.; Raychowdhury, R.; Zeng, Q.; et al. Full-length transcriptome assembly from RNA-Seq data without a reference genome. *Nat. Biotechnol.* **2011**, *29*, 644–652. [[CrossRef](#)] [[PubMed](#)]
122. Paytavi Gallart, A.; Hermoso Pulido, A.; Anzar Martinez de Lagrán, I.; Sanseverino, W.; Aiese Cigliano, R. GREENC: A Wiki-based database of plant lncRNAs. *Nucleic Acids Res.* **2015**, *44*, D1161–D1166. [[CrossRef](#)] [[PubMed](#)]
123. Shimomura, M.; Kanamori, H.; Komatsu, S.; Namiki, N.; Mukai, Y.; Kurita, K.; Kamatsuki, K.; Ikawa, H.; Yano, R.; Ishimoto, M.; et al. The *Glycine max* cv. Enrei genome for improvement of Japanese soybean cultivars. *Int. J. Genom.* **2015**, *2015*, 358127.
124. Duvick, J.; Fu, A.; Muppirala, U.; Sabharwal, M.; Wilkerson, M.D.; Lawrence, C.J.; Lushbough, C.; Brendel, V. PlantGDB: A resource for comparative plant genomics. *Nucleic Acids Res.* **2007**, *36*, D959–D965. [[CrossRef](#)] [[PubMed](#)]
125. Goodstein, D.M.; Shu, S.; Howson, R.; Neupane, R.; Hayes, R.D.; Fazo, J.; Mitros, T.; Dirks, W.; Hellsten, U.; Putnam, N.; et al. Phytozome: A comparative platform for green plant genomics. *Nucleic Acids Res.* **2011**, *40*, D1178–D1186. [[CrossRef](#)] [[PubMed](#)]



© 2018 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0/>).