Correlation between Metabolic Specialization and Codon Preference: Analysis of the Ligninolytic Genes from the White Rot Basidiomycete Ceriporiopsis Subvermispora as A Model System.

Alex R Gonzalez
Universidad de Los Lagos
Gino R Corsini
Universidad Autonoma
Sergio Lobos
Universidad de Chile
Daniela Seelenfreund
Universidad de Chile
Mario Tello (mario.tello@usach.cl)
Universidad de Santiago de Chile https://orcid.org/0000-0003-4573-6460

Research article

Keywords:

Posted Date: June 9th, 2020

DOI: https://doi.org/10.21203/rs.3.rs-32616/v1

Abstract

Background: *Ceriporiopsis subvermispora* is a white-rot fungus that displays a high specificity towards lignin mineralization when colonizing dead wood or lignocellulosic compounds. The lignocellulose degrading system from *C. subvermispora* is formed by genes that encode cellulose hydrolytic enzymes, manganese peroxidases, and laccases that catalyze the efficient depolymerization and mineralization of lignin in the presence of Mn^{3+} through the formation of lipoperoxides from unsaturated lipid acids. This highly specific lignin-degrading system is unique among white-rot fungi.

Methods: In order to determine if this metabolic specialization has modified codon usage of the ligninolytic system, leading to an increased adaptation to the fungal translational machine, we analyzed the adaptation to host codon usage (CAI), tRNA pool (tAI, and AAtAI), codon pair bias (CPB) and the number of effective codons (Nc). These indexes were correlated with gene expression of *C. subvermispora,* as evaluated by microarray in the presence of two carbon sources, glucose and Aspen wood.

Results: General gene expression of *C. subvermispora* was not correlated with the CAI, tAI, AAtAI, CBP or Nc indexes used to evaluate adaptation to codon bias or the tRNA pool, neither in the presence of glucose or Aspen wood. However, in media containing Aspen wood, the induction of expression of lignin-degrading genes showed a strong correlation with all the former indexes. Lignin-degrading genes, defined as genes whose expression increases at least two-fold in Aspen wood, showed significantly ($p<0.001$) higher values of CAI, AAtAI, CPB, tAI and lower values of Nc with respect to non-induced genes. Among ligninolytic genes, cellulose-binding proteins and manganese peroxidases presented the highest adaptation values. We also identified an expansion of genes encoding glycine and glutamic acid tRNAs.

Conclusions: Our results suggest that the metabolic specialization to use wood as the sole carbon source has introduced a bias in the codon usage of genes involved in lignocellulose degradation. This bias reduces codon diversity and increases codon usage adaptation to the tRNA pool available in *C. subvermispora*. To our knowledge, this is the first study showing that codon usage is modified to improve the translation efficiency of a group of genes involved in a particular metabolic pathway.

1.- Introduction

The main carbon source synthesized through photosynthesis which plays a central role in the carbon cycle of the planet is lignocellulose. Lignin, one of the compounds of lignocellulose, is a recalcitrant, aromatic, and amorphous polymer that protects lignocellulose from microbial attack. A small group of filamentous fungi from the basidiomycete phylum are unique in their ability to efficiently degrade lignin [1]. Together they are collectively known as white rot fungi, developing an enzymatic machinery that allows degradation of the three main components of lignocellulose: lignin, cellulose and hemicellulose. White-rot fungi mineralize lignin as a strategy to access cellulose and hemicellulose, whose sugar moieties are used as carbon and energy sources. Degradation of these aromatic and carbohydrate polymers progresses by different mechanisms. While mineralization of lignin is carried out by free radicals generated enzymatically [2, 3], the degradation of cellulose and hemicellulose into its constituent sugars occurs through the combination of direct enzymatic hydrolysis and partial hydrolysis by enzymatically generated free radicals [4].

Ceriporiopsis subvermispora is a white-rot basidiomycete that mineralizes lignin using a machinery composed of manganese peroxidase, versatile peroxidase, and laccase, as well as a system generating lipoperoxides [5–9]. Biochemical and genetic evidence shows that manganese peroxidases and laccases are encoded by a family of structurally related genes [6, 8, 10–12]. Hydrolysis of cellulose and hemicellulose is also achieved by a suite of cellulases and hemicellulases, which are encoded by multigene families [8]. Gene expression of enzymes involved in lignocellulose digestion and mineralization is induced by compounds present in wood, as shown by research using Northern blot, RT-PCR and microarrays [8, 13–17]. These studies also indicate that expression levels differ among members of a given multigene family [8, 11, 18]. Although selective pressure imposed by lignocellulose metabolism is reflected by an increase in the copy number of ligninolytic genes in *C. subvermispora* and other white-rot fungi [8], it is unknown whether this condition has influenced the composition of synonymous codons in these genes.

Preference for use of some synonymous codons (codon bias) is influenced by two mechanisms, mutational bias, and translational selection. Mutational bias establishes that the codon preference of a gene is determined by the GC content of the organism or the region where the gene is located in the genome [19]. On the other hand, translational selection establishes that codon bias is determined by the influence of synonymous mutations on the translational process [20, 21]. Highly expressed genes, such as those coding for ribosomal proteins, tend to have a greater bias in the use of synonymous codons, preferentially using codons which are most represented in the genome [22]. In model organisms such as *E. coli* and *S. cerevisiae*, tRNAs that decode frequently used codons tend to have a larger gene copy number [23, 24]. Thus, genes with frequently used codons are more adapted to the tRNA pool and consequently have more tRNAs available for its translation, which in turn improves the efficiency of protein synthesis.

In this work, we compared codon usage, adaptation to the tRNA pool, and codon bias of *C. subvermispora* genes involved in lignocellulose metabolism and genes related to other cell processes. Our results show that selective pressure imposed by the use of lignocellulose has specifically modified codon usage of genes involved in the utilization of lignocellulose, favoring an increase in translational efficiency with respect to genes not involved in this process.

2.- Materials Y Methods

Sequences. Genome sequences from *C. subvermispora* were downloaded from https://mycocosm.jgi.doe.gov/Cersu1/Cersu1.home.html

Microarray data. Microarray data were downloaded from the Gene Expression Omnibus Database, accession no. GSE34636 [8].

Determination of the gene copy number coding for tRNAs. Genes encoding tRNAs were identified using the tRNAscan -SE program, using unmasked scaffolds assembled from the genome sequence of *C. subvermispora* [8, 25].

Determination of codons and codon pair frequencies. The frequency of codon usage and codon pairs was determined from the coding sequences of *C. subvermispora*. To calculate the frequency of codon usage the program JEMBOSS through the CUSP routine was used (http://emboss.sourceforge.net/Jemboss/) [26]. A written script in Python language was used for determining the frequency of codon pairs [27].

Determination of bias in codon-pair usage. Bias in codon-pair usage was determined as described by Coleman [28]. The observed frequency of amino acid pairs was deduced based on the sum of all the codon pairs that codify for the same combination of amino acids. Subsequently, we used these frequencies to calculate the CPS index (codon-pair score) using the equation described by Mueller *et al*. 2006 [29] (Eq. 1), where f(AB) corresponds to the frequency of the appearance of the AB codon pair expressed in parts-per-thousands with respect to the total number of codon pairs in the genome. f(A) and f(B) correspond to the abundance of each of the individual codons expressed in parts-per-thousands with respect to the total number of codons in the genome. f(X) and f(Y) correspond to the amino acids codified by the A and B codons, respectively, and f(XY) is the abundance of the amino acid pairs in all the proteins of the organism under study. The CPS values of the 3271 codon pairs in *C. subvermispora* can be found in the supplementary material (**Additional file 1**). We calculated the bias in codon usage as the arithmetic mean of the codon-pair score of a gene (CPS) as described by Coleman *et al*. (2006) (Eq. 2), in which i is the i^{th} codon pair of a gene and 1 is the total length of the gene expressed in the quantity of codified amino acids.

$$CPS = Ln\left(\frac{f(AB)}{f(A)\times f(B)} \times \frac{f(X)\times f(Y)}{f(XY)}\right)$$

Equation 1

$$CPB = \frac{\sum_i^l CPS_i}{l-1}$$

Equation 2

Calculation of RSCU and CAI. The Relative Synonymous Codon Usage (RSCU) and Codon Adaptation Index (CAI) were calculated with the Emboss program [30] based on the frequency of codon usage of *C. subvermispora*.

Determination of adaptation to the tRNA pool. The adaptation to the tRNA pool of genes present in *C. subvermispora* was determined by calculating the values of tAI and AAtAI [31]. We calculated tAI as established by dos Reis [24] (equations 3 and 4), estimating the relative abundance of tRNAs from the number of copies of each tRNA gene in the *C. subvermispora* genome. The number of copies of each tRNA gene was determined with the tRNAscan-SE program [25] in Linux, which was used to analyze the unmasked assembled *C. subvermispora* genome.

Equation 3

W_i is the relative adaptiveness of the i^{th} codon to the tRNA pool, n_i is the number of tRNA isoacceptors that recognize the i^{th} codon, $tGCN_{ij}$ is the number of copies of the j^{th} tRNA gene that recognizes the i^{th} codon, and S_{ij} is the selective constraint in the efficiency of codon-anticodon pairing.

Equation 4

The adaptation of a gene to the tRNA pool is calculated according to Eq. 4, in which w_i is defined as the ratio between W_i and W_{max} (W_i/W_{max}), ikg is the codon defined by the k^{th} triplet of gene g, and l_g is the length of gene g in codifying codons.

Determination of AAtAI. Briefly, AAtAI was calculated using Eq. 5.

Equation 5

$^{AA}w_i$ is defined as the ratio between W_i and W_{AAmax} (W_i/W_{AAmax}), where W_{AAmax} is the highest value of Wi among codons coding for the same aminoacid. ikg is the codon defined by the k^{th} triplet of gene g, and l_g is the length of gene g in codifying codons. As AAtAI is calculated from Eq. 5, which is similar to the calculation of CAI, we used the EMBOSS software [30], entering Wi data in replacement of the frequency of codon usage. Wi was calculated using the procedure described for tAI in Eq. 3.

Determination of Nc: The number of effective codons for the *C. subvermispora* genes was calculated using the CodonW program (http://codonw.sourceforge.net/)

Phylogenetic Analysis

Multialignment between sequences of tRNA genes was performed using ClustalW [32]. The parameters were set up to align sequences using IUB as substitution matrix. The evolutionary history was inferred using the Neighbor-Joining method [33]. Interior branch test with a bootstrap of 1000 was used to analyze confidence of the tree [34]. The evolutionary distances were computed using the Maximum Composite Likelihood method [35]. The rate of variation among sites was modeled with a gamma distribution (shape parameter = 1). Multialignment and evolutionary analyses were conducted in MEGA5 [36].

Graphs and statistical methods. The program SigmaPlot 11 was used for graphs and statistical tests. The significance of the differences or correlations among the data groups obtained was evaluated with the Rank Sum Test non-parametric test for comparing two groups and the non-parametric Spearman Rank Order test for correlations, using a value of $p < 0.05$ as a cutoff.

3.- Results

C. subvermispora **tRNAs**. Genome analysis of *C. subvermispora* by tRNAScan-SE identified a total of 192 tRNAs in 32 scaffolds (Fig. 1). About 72% of the tRNA genes presented introns (Table 1). The scaffold with the highest number of tRNA genes was scaffold 1, which contains 20 copies of various tRNA genes (**Additional file 2**). The tRNA with the highest number of gene copies corresponded to tRNAs charging glycine, with 17 gene copies distributed in eight scaffolds (scaffolds 1, 2, 5, 7, 9, 10, 13, and 20). The tRNAs for cysteine and tryptophan presented the lowest number of gene copies, each with three gene copies in three scaffolds (Table 1).

Phylogenetic analysis of C. subvermispora **tRNA genes**. To determine whether the high number of tRNA genes charging the same amino acid corresponds to related gene copies, a phylogenetic reconstruction and evolutionary distance calculation were performed using the tRNA sequences identified by the tRNAscan-SE program. Phylogenetic reconstruction indicated that most tRNA genes that code for the same amino acid group together, with the exception of the tRNA genes that charge arginine, valine, and alanine (Fig. 1), which form two groups in each case. For tRNAs loading arginine, Group I comprise genes presenting anticodons with the WCG sequence. In contrast, group II presents a YCK anticodon consensus sequence. In tRNAs that load valine, Group I correspond to tRNA genes without introns, while group II includes all valine tRNA genes containing introns. Genes coding for the group I of alanine tRNAs exhibit anticodons with the consensus sequence YGC, whereas, the anticodon sequence is AGC for group II (Fig. 1). The tRNA genes 66 and 155 also showed a different pattern. tRNAScan prediction indicates that the amino acid loaded by tRNA155 should be serine, however, the sequence of this gene grouped with threonine charging tRNA genes. Also, the tRNA66 gene is expected to load isoleucine though this gene does not group with isoleucine charging tRNAs (Fig. 1). To identify tRNA genes that are repeated, the evolutionary distance between different tRNA genes was calculated. The genes with values of evolutionary distance equal to 0.000 were selected. This analysis identified 15 tRNA genes that are repeated between two and ten times. The group of tRNA genes that showed the greatest expansion corresponds to those tRNAs carrying glutamic acid and glycine. One glycine tRNA genes is repeated twice and the other is repeated ten times. The tRNA gene for glutamic acid is also repeated ten times (**Additional file 3**). These genes are scattered along the genome, with the exception of tRNA genes 91, 92, 94, 95, 96, and 97, which code for glycine and are adjacent in the genome.

tRNA abundance and codon usage in C. subvermispora. The expansion of certain families of tRNAs in the genome of *C. subvermispora* could be the result of an evolutionary pressure to increase its expression. In some organisms such as *E. coli* and yeast, the number of copies of a tRNA gene is proportional to the abundance in the genome of the decoded codon [24]. This proportionality is explained because during translation process, tRNAs that decode frequently used codons have a higher rate of consumption. To sustain adequate translation, cells must balance synthesis and consumption rates of tRNAs. The increase in copy number of a gene that encodes a tRNA that recognizes a frequently used codon, allows increasing expression of this tRNA to balance its rate of consumption.

To test if the expansion of certain families of tRNAs in *C. subvermispora* is related to the increment of the specific codons, we assessed whether there was a correlation between the frequency of codon usage and the amount of tRNA genes that decode these most highly used codons. We identify a positive correlation between these parameters (**$\rho = 0.406$, $p = 0.0016$, $n = 61$).** Dos Reis et al show that the Relative Adaptativeness to the tRNA pool (w) which takes in account that codons can be recognized by anticodons with perfect or imperfect match (wobble codon-anticodon recognition rules) with different affinities is a better parameter to measure the adaptation of a codon to their decoding tRNAs than the absolute number of tRNA. When the frequency of codon usage was correlated with the Relative Adaptativeness to the tRNA pool (w), an increased correlation was observed (**$\rho = 0.459$, $p = 2.2 \times 10^{-4}$, $n = 61$**) (Fig. 2**)**.

Synonymous codons are not used equally in an organism. As one tRNA can decode several synonymous codons with different affinities, the expansion of some tRNA families in *C. subvermispora* may be related to the preferential use of certain synonymous codons in coding regions of *C. subvermispora*. To assess this hypothesis, we correlated RSCU values with Relative Adaptativeness to the tRNA pool. Non-statistical correlation was observed, in part because w values are normalized with respect to the tRNA with the highest number of genes able to decode it (W_i/W_{max}) and not with respect to the pool of tRNAs that decode the complete set of synonymous codons. When W values where normalized with respect to the total amount of tRNAs that decode a set of synonymous codons, a strong correlation with the RSCU values (**$\rho = 0.628$, $p = 0$, $n = 61$**) was observed. This suggests that among synonymous codon, those highly represented in the *C. subvermispora* genome tend to be decoded by those tRNAs with a high number of gene copies.

Relationship between gene expression level, codon bias, and translational efficiency in C. subvermispora. The bias in codon usage and adaptation to the tRNA pool modulates translational efficiency. Thus, highly expressed genes tend to use codons that are over-represented in the genome, which in turn present greater availability of tRNAs [24]. To determine whether this relationship exists in *C. subvermispora*, expression levels of *C. subvermispora* genes were correlated with their adaptation values to the tRNA pool and to codon bias. Adaptation to the tRNA pool was evaluated using two indexes: i) tAI, which measures adaptation of the tRNA pool compared to the relative amount of each tRNA gene, and ii) AAtAI, which evaluates whether a gene preferentially uses the most abundant tRNA charging a particular amino acid. Codon bias was analyzed by calculation of CAI (Codon Adaptation Index), and CPB (Codon Pair-Bias), to assess if gene expression is correlated with bias in usage of codons or of codon pairs. Codon bias also was evaluated using Nc value to determine if the increase in transcription is associated with a decrease in the diversity of codons used. We analyzed expression levels determined by RNAseq obtained from the fungus grown on a medium with Ball-Milled Aspen [37]. All indicators showed high degrees of correlation among them, however, these indicators show very low correlation coefficients with expression levels measured by RNAseq (**Additional file 4**).

Transcriptional response to growth on BMA, codon bias and translational efficiency. Growth of *C. subvermispora* in natural environments is dependent on wood. In 2012, Fernandez-Fueyo et al reported microarray experiments that compared gene expression of *C. subvermispora* grown on glucose and on Ball-Milled Aspen (BMA) as carbon sources. Saline media with BMA has been used as a laboratory medium that mimics growth on wood to analyze expression of genes that are transcriptionally regulated by growth on wood. To analyze if genes regulated by conditions that mimic growth on wood, such as BMA have a different adaptation to the tRNA pool or codon bias, we used the microarray data published by Fernandez-Fueyo and defined four groups of genes: group A corresponds to genes where expression was reduced at least 2 times with a p-value lower than 0.05. Group B includes all genes which showed increased

expression of at least 2 times with p-value lower than 0.05. Group C corresponds to all genes with non-significant differences ($p > 0.05$) and group D contains all genes with low changes in expression (< 2 fold) that are statistically significant. ($p < 0.05$). Our results show that group B has lower values of Nc and higher values of CAI, tAI, AAtAI, and CPB. Groups A, C, and D show non-significant differences among them. This implies that genes induced by wood preferentially use a reduced set of codons that are better adapted to the tRNA pool present in *C. subvermispora* (Fig. 3).

When we correlated the CAI, tAI, AAtAI, Nc and CPB indexes with the ratio between expression in BMA and glucose culture medium, a statistically significant correlation was observed. Positive correlations were found with almost all indicators used (CAI, tAI, AAtAI and CPB), the exception was Nc that showed a negative correlation. The higher correlation were identified in CAI and tAI indexes, and in genes that showed significant differences for expression in BMA saline medium compared to expression in glucose-supplemented saline medium (Table 3). This increase in correlation coefficients can be explained if growth on lignocellulose exerts pressure on codon usage of genes involved in the metabolization of this carbon source, thereby selecting those codons that increase the translational efficiency of these genes.

Interestingly, when genes from Group B were sorted according to their codon usage adaptation values or to the tRNA pool, we found that genes coding for ribosomal proteins presented the highest CAI, tAI, and AAtAI values **(Additional file 5)**. An increase in the expression of these genes may lead to an increase of the protein biosynthetic capacity, thus exposure of *C. subvermispora* to wood or lignocellulose leads to an increase of the rate of translation, specifically enhancing the synthesis of ligninolytic enzymes.

Translational efficiency and codon bias in ligninolytic genes. Genome analysis of *C. subvermispora* indicates that this organism has an expansion of the number of genes directly related to the mineralization and hydrolysis of lignocellulose. The genome contains 17 annotated genes of manganese peroxidase, five genes coding for laccases, four genes for cellobiose dehydrogenases (CDH), six genes for Δ-12 dehydrogenases, and five genes for Δ-9 dehydrogenases. These genes that encode the enzymes that catalyze the same reaction exhibit differential expression, which might reflect that they serve different functions in the mineralization/hydrolysis process of lignocellulose [8]. To determine whether genes of the same family present a similar bias of codon usage and of codon adaptation or adaptation to the tRNA pool, we compared the values of Nc, tAI, AAtAI, CAI, and CPB of these genes. These values were also normalized respect to the mean and standard deviation of the values obtained for all genes encoded in *C. subvermispora* (Z-values) [38]. This normalization was applied to identify if and how values for ligninolytic genes differ with respect to these same values in genes not directly related to the ligninolytic process (**Table 4**).

Genes encoding manganese peroxidase generally show above-average adaptation values to the tRNA pool and codon usage (Z value > 0). The manganese peroxidase gene with the highest level of expression (ID: 50297) proved to be the most adapted to the tRNA pool, with tAI values that are more than two standard deviations from the mean tAI values of other *C. subvermispora* genes (Z-tAI > 2). We also observed that the most induced manganese peroxidase gene in BMA medium also shows a high value of adaptation to the tRNA pool with a (Z-tAI = 1.619). Genes encoding laccases showed a similar trend, as only the gene that significantly changed their expression levels after growth in BMA medium (ID130783) showed a value of tAI above the average (Z-tAI = 0.461).

Among cellulases, the gene ID148588 showed the highest adaptation values to the tRNA pool (Z-tAI = 1.756) and high expression level under growth with glucose as the sole carbon source. This gene also increased its expression in BMA medium. Several genes for cellulases show Nc values below 40, indicating a strong bias in the use of synonymous codons. However, in this group of genes an association between expression levels or adaptation to the tRNA pool or any other index used in this work was not found. Genes for CDH show little bias in codon usage (Nc ~ 50). Only the CDH gene which is induced by BMA (ID: 84792) shows an Nc value of 48 and values of adaptation to the tRNA pool slightly higher than mean (Z > 0) (**Table 4**). The Δ12-dehydrogenase genes showed a weak codon bias with the exception of gene ID124050, which showed a Nc value of 38 and an above average adaptation to the tRNA pool. This gene also showed a strong expression in medium containing glucose which was not modified by the BMA medium. An opposite behavior was observed in the 9-dehydrogenase genes, where the gene with the greatest induction in BMA (ID129048) showed a lower Nc value (41.46), suggesting a strong bias in codon usage (**Table 4**).

When genes related to mineralization and digestion of lignocellulose were ordered according to their adaptation values to codon usage or the tRNA pool, genes with greater adaptation values were found to belong to the manganese peroxidase families and to proteins with cellulose-binding domains. Interestingly, manganese peroxidases are more adapted to the tRNA pool, while proteins with cellulose-binding domains showed higher adaptation to codon usage.

4.- Discussion

The development of massive sequencing technologies, bioinformatics sequence analysis and synthetic biology have established that synonymous mutations, far from being silent, play an important role in the fine-tuning of protein synthesis efficiency and the role of different functional forms [39, 40]. The first clues or hints of the importance of synonymous mutations for translational efficiency arose from the identification of bias in the use of synonymous codons that were detected in highly expressed ribosomal proteins from *E. coli*, *Bacillus subtilis* and *S. cerevisiae* [22, 41, 42]. In viruses, genes coding for highly required proteins, such as those of the virus capsid or nucleoprotein, show higher adaptation values to codon usage of the host than other viral genes [38, 43]. The generation of synthetic genes has revealed that the amount of synthesized protein can vary over thousand fold solely by changing the composition of synonymous codons [44]. Synthetic viruses, constructed from highly virulent viruses, show reduced replication and very low mortality when their codon usage is changed to codon combinations present in low frequency in the host genome [45]. Further, the arrangement of codons within a gene is not random and is related to the proper folding of the proteins and also to the proper recycling of ribosomes [31]. For example, the low adaptation level of the FRQ gene of *Neurospora crassa* to codon usage is essential to maintain the rhythm of the circadian clock [46]. Continuously increasing evidence supports the fact that bias in the use of synonymous codons plays an active role in the fine-tuning of protein production. The relationship between the use of synonymous codons and translation efficiency lies in the abundance of cognate tRNAs. Highly used codons tend to have a greater number of tRNAs that recognize them [47]. In this work, we have identified 192 genes encoding tRNAs in the white-rot basidiomycete *C. subvermispora*, a number that is similar to the quantity of tRNA genes present in other

fungi such as *Aspergillus fumigatus* (178 genes) and *Schizosaccharomyces pombe* (186 genes) (http://gtrnadb.ucsc.edu/**).** Moreover, *C. subvermispora* tRNA genes present the interesting feature that each type of tRNA gene (i. e. that loads the same amino acid), groups in a different clade, indicating that the current pool of tRNAs present in the *C. subvermispora* monokaryotic strain B has arisen from gene duplication processes, and probably not from horizontal transfer or recombination. This result is consistent with the absence of a known *C. subvermispora* sexual stage [12]. Interestingly, we identified an expansion of the tRNA genes coding for amino acids glycine and glutamine. The high similarity of their sequences indicates that this expansion may have occurred recently. The expansion of this set of tRNA genes could be the result of the presence of some unidentified SINE elements, which are scored as tRNAs by the tRNAscan program [48]. Another alternative is that this expansion is guided by the need to synthesize large quantities of glycine- and glutamic acid-rich proteins, requiring this set of tRNA genes. This strategy is similar to that used by bacteriophages to increase their rate of protein synthesis in a new host, that involves carrying tRNA genes for those codons that are most used in bacteriophage genes [49]. Recently it was discovered that HIV also uses a similar strategy, specifically packaging a pool of tRNAs whose codons are poorly represented in the human genome [50]. Interestingly, the addition of glutamic acid to the culture medium increases the production of cellulases by the brown rot fungus *Fomitopsis* sp. RCK2010 [51], and increases the production of manganese peroxidases and laccases in some white rot fungi [52]. On the other hand, the amino acid glycine is a precursor of heme biosynthesis [53], an important cofactor present in manganese peroxidases and cytochromes, genes that are highly abundant in *C. subvermispora*.

We also found a strong bias in codon usage, codon pairs, and adaptation to the tRNA pool of *C. subvermispora* genes involved in lignocellulose degradation. Similar bias has been described in Lignin peroxidase from *Phaenorochaete chrysosporium*[54]. Since a bioinformatic identification of these genes can only be applied to those directly involved in the mineralization/digestion of lignocellulose, such as manganese peroxidase type enzymes, laccases and cellulases, we devised a functional definition, whereby genes involved in lignocellulose degradation are those induced in the presence of BMA, a model substrate of wood. Studies of phylogenetic reconstruction indicate that the lignocellulose-degrading machinery may have arisen during the Paleozoic [55]. The efficient use of this carbon source requires a metabolic adaptation that involves the whole organism, rather than the emergence of new types of enzymes. Therefore, it is expected that enzymes that are part of the general metabolism of *C. subvermispora* also respond to exposure of lignocellulose, and that they become necessary for the efficient use of this carbon source. Expansion of cellulase, laccase, and manganese peroxidase gene families in the *C. subvermispora* genome permitted that each of these genes developed differential expression; however, the precise contribution of each of these genes in the digestion and mineralization of lignocellulose is not clear. Using this functional definition, we were able to detect biases for all evaluated parameters. In turn, the lack of strong correlation between gene expression level and adaptation to the tRNA pool or codon bias of *C. subvermispora* grown in a medium containing glucose or BMA, is consistent with this postulate.

In bacteria, an increase of copy numbers of a particular gene is a strategy for increasing its expression. This genomic adaptation could explain the expansion of genes directly related to the digestion and mineralization of lignocellulose. However, the complex expression pattern of these genes suggests that this gene expansion process is more intricate than simply a response to the need of producing more enzymes. The high bias in codon usage, adaptation of the tRNA pool and codon diversity, in some genes directly related to the digestion and mineralization processes of lignocellulose together with the increase in expression of ribosomal proteins in BMA suggest that *C. subvermispora* also uses increased translational efficiency as an additional strategy to increase the production these specific set of proteins. Thus, the increase in copy number may be linked to the generation of a diverse array of enzymes that can process a wide range of substrates. In support of this hypothesis, it has been shown that *C. subvermipora* manganese peroxidases present different kinetic parameters and substrate specificities [8, 9].

5. - Conclusions

Our results suggest that lignocellulose degradation by *C. subvermispora* has modified the genome structure of this fungus, changing the bias in codon usage, the tRNA gene pool, and codon diversity in genes that are induced in the presence of wood substrates, in order to optimize the production of these proteins. This strategy may be particularly useful in slow-growing organisms, such as *C. subvermispora* that cannot increase the production of enzymes by increasing cell mass. To our knowledge, this study is the first example to show metabolic adaptation to a particular ecological niche by modification of the genetic structure of an organism in favor of a selective increase of the translational efficiency of genes involved in metabolizing specific substrates that determine its adaptation to a particular environment.

10.- Declarations

10.1.- Ethics approval and consent to participate

Not applicable

10.2.- Consent for publication

Not applicable

10.3.- Competing interests

The authors declare that they have no competing interests

10.4.- Funding.

021971GM_DAS to MT

10.5.- Authors' contributions

MT Design the in silico and bioinformatic experiments.

AG calculated tAI and AAtAI values

GC calculated CAI and CPB values

SL studied the relation between tAI, AAtAI, CAI and CPB with the gene expression.

DS helped to analyze the evolutive implication of bias in codon and adaptation to the tRNA pool.

10.6.- Acknowledgements

The author acknowledge to Dan Cullen for the critical review of this manuscript.

6.- References

1. Martínez AT, Speranza M, Ruiz-Dueñas FJ, Ferreira P, Camarero S, Guillén F, et al. Biodegradation of lignocellulosics: microbial, chemical, and enzymatic aspects of the fungal attack of lignin. Int Microbiol. 2005;8:195–204.
2. Kirk TK, Farrell RL. Enzymatic "combustion": the microbial degradation of lignin. Annu Rev Microbiol. 1987;41:465–505. doi:10.1146/annurev.mi.41.100187.002341.
3. Kersten P, Cullen D. Extracellular oxidative systems of the lignin-degrading Basidiomycete Phanerochaete chrysosporium. Fungal Genet Biol. 2007;44:77–87. doi:10.1016/j.fgb.2006.07.007.
4. Baldrian P, Valásková V. Degradation of cellulose by basidiomycetous fungi. FEMS Microbiol Rev. 2008;32:501–21. doi:10.1111/j.1574-6976.2008.00106.x.
5. Rüttimann-Johnson C, Salas L, Vicuña R, Kirk TK. Extracellular Enzyme Production and Synthetic Lignin Mineralization by Ceriporiopsis subvermispora. Appl Environ Microbiol. 1993;59:1792–7.
6. Lobos S, Larraín J, Salas L, Cullen D, Vicuña R. Isoenzymes of manganese-dependent peroxidase and laccase produced by the lignin-degrading basidiomycete Ceriporiopsis subvermispora. Microbiology. 1994;140(Pt 1):2691–8.
7. Enoki M, Watanabe T, Nakagame S, Koller K, Messner K, Honda Y, et al. Extracellular lipid peroxidation of selective white-rot fungus, Ceriporiopsis subvermispora. FEMS Microbiol Lett. 1999;180:205–11.
8. Fernandez-Fueyo E, Ruiz-Dueñas FJ, Ferreira P, Floudas D, Hibbett DS, Canessa P, et al. Comparative genomics of Ceriporiopsis subvermispora and Phanerochaete chrysosporium provide insight into selective ligninolysis. Proc Natl Acad Sci U S A. 2012;109:5458–63. doi:10.1073/pnas.1119912109.
9. Fernández-Fueyo E, Ruiz-Dueñas FJ, Miki Y, Martínez MJ, Hammel KE, Martínez AT. Lignin-degrading peroxidases from genome of selective ligninolytic fungus Ceriporiopsis subvermispora. J Biol Chem. 2012;287:16903–16. doi:10.1074/jbc.M112.356378.
10. Salas C, Lobos S, Larraín J, Salas L, Cullen D, Vicuña R. Properties of laccase isoenzymes produced by the basidiomycete Ceriporiopsis subvermispora. Biotechnol Appl Biochem. 1995;21(Pt 3):323–33.
11. Tello M, Corsini G, Larrondo LF, Salas L, Lobos S, Vicuña R. Characterization of three new manganese peroxidase genes from the ligninolytic basidiomycete Ceriporiopsis subvermispora. Biochim Biophys Acta. 2000;1490:137–44.
12. Tello M, Seelenfreund D, Lobos S, Gaskell J, Cullen D, Vicuña R. Isolation and characterization of homokaryotic strains from the ligninolytic basidiomycete Ceriporiopsis subvermispora. FEMS Microbiol Lett. 2001;199:91–6.
13. Manubens A, Canessa P, Folch C, Avila M, Salas L, Vicuña R. Manganese affects the production of laccase in the basidiomycete Ceriporiopsis subvermispora. FEMS Microbiol Lett. 2007;275:139–45. doi:10.1111/j.1574-6968.2007.00874.x.
14. Gutiérrez M, Rojas LA, Mancilla-Villalobos R, Seelenfreund D, Vicuña R, Lobos S. Analysis of manganese-regulated gene expression in the ligninolytic basidiomycete Ceriporiopsis subvermispora. Curr Genet. 2008;54:163–73. doi:10.1007/s00294-008-0209-7.
15. Alvarez JM, Canessa P, Mancilla RA, Polanco R, Santibáñez PA, Vicuña R. Expression of genes encoding laccase and manganese-dependent peroxidase in the fungus Ceriporiopsis subvermispora is mediated by an ACE1-like copper-fist transcription factor. Fungal Genet Biol. 2009;46:104–11. doi:10.1016/j.fgb.2008.10.002.
16. Mancilla RA, Canessa P, Manubens A, Vicuña R. Effect of manganese on the secretion of manganese-peroxidase by the basidiomycete Ceriporiopsis subvermispora. Fungal Genet Biol. 2010;47:656–61. doi:10.1016/j.fgb.2010.04.003.
17. Hori C, Gaskell J, Igarashi K, Kersten P, Mozuch M, Samejima M, et al. Temporal alterations in the secretome of the selective ligninolytic fungus Ceriporiopsis subvermispora during growth on aspen wood reveal this organism's strategy for degrading lignocellulose. Appl Environ Microbiol. 2014;80:2062–70. doi:10.1128/AEM.03652-13.
18. Manubens A, Avila M, Canessa P, Vicuña R. Differential regulation of genes encoding manganese peroxidase (MnP) in the basidiomycete Ceriporiopsis subvermispora. Curr Genet. 2003;43:433–8. doi:10.1007/s00294-003-0410-7.
19. Sharp PM, Stenico M, Peden JF, Lloyd AT. Codon usage: mutational bias, translational selection, or both? Biochem Soc Trans. 1993;21:835–41.
20. Supek F, Skunca N, Repar J, Vlahovicek K, Smuc T. Translational selection is ubiquitous in prokaryotes. PLoS Genet. 2010;6:e1001004. doi:10.1371/journal.pgen.1001004.
21. Tuller T, Waldman YY, Kupiec M, Ruppin E. Translation efficiency is determined by both codon bias and folding energy. Proc Natl Acad Sci U S A. 2010;107:3645–50. doi:10.1073/pnas.0909910107.

22. Sharp PM, Li WH. The codon Adaptation Index–a measure of directional synonymous codon usage bias, and its potential applications. Nucleic Acids Res. 1987;15:1281–95.
23. Bulmer M. Coevolution of codon usage and transfer RNA abundance. Nature. 1987;325:728–30. doi:10.1038/325728a0.
24. dos Reis M, Savva R, Wernisch L. Solving the riddle of codon usage preferences: a test for translational selection. Nucleic Acids Res. 2004;32:5036–44. doi:10.1093/nar/gkh834.
25. Lowe TM, Eddy SR. tRNAscan-SE: a program for improved detection of transfer RNA genes in genomic sequence. Nucleic Acids Res. 1997;25:955–64.
26. Carver T, Bleasby A. The design of Jemboss: a graphical user interface to EMBOSS. Bioinformatics. 2003;19:1837–43.
27. Tello M, Saavedra JM, Spencer E. Analysis of the use of codon pairs in the HE gene of the ISA virus shows a correlation between bias in HPR codon-pair use and mortality rates caused by the virus. Virol J. 2013;10:180. doi:10.1186/1743-422X-10-180.
28. Coleman JR, Papamichail D, Skiena S, Futcher B, Wimmer E, Mueller S. Virus attenuation by genome-scale changes in codon pair bias. Science. 2008;320:1784–7. doi:10.1126/science.1155761.
29. Mueller S, Papamichail D, Coleman JR, Skiena S, Wimmer E. Reduction of the rate of poliovirus protein synthesis through large-scale codon deoptimization causes attenuation of viral virulence by lowering specific infectivity. J Virol. 2006;80:9687–96. doi:10.1128/JVI.00738-06.
30. Rice P, Longden I, Bleasby A. EMBOSS: the European Molecular Biology Open Software Suite. Trends Genet. 2000;16:276–7.
31. Tuller T, Carmi A, Vestsigian K, Navon S, Dorfan Y, Zaborske J, et al. An evolutionarily conserved mechanism for controlling the efficiency of protein translation. Cell. 2010;141:344–54. doi:10.1016/j.cell.2010.03.031.
32. Thompson JD, Higgins DG, Gibson TJ. CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. Nucleic Acids Res. 1994;22:4673–80.
33. Saitou N, Nei M. The neighbor-joining method: a new method for reconstructing phylogenetic trees. Mol Biol Evol. 1987;4:406–25.
34. Dopazo J. Estimating errors and confidence intervals for branch lengths in phylogenetic trees by a bootstrap approach. J Mol Evol. 1994;38:300–4.
35. Tamura K, Nei M, Kumar S. Prospects for inferring very large phylogenies by using the neighbor-joining method. Proc Natl Acad Sci U S A. 2004;101:11030–5. doi:10.1073/pnas.0404206101.
36. Tamura K, Peterson D, Peterson N, Stecher G, Nei M, Kumar S. MEGA5: molecular evolutionary genetics analysis using maximum likelihood, evolutionary distance, and maximum parsimony methods. Mol Biol Evol. 2011;28:2731–9. doi:10.1093/molbev/msr121.
37. Hori C, Gaskell J, Igarashi K, Kersten P, Mozuch M, Samejima M, et al. Temporal alterations in the secretome of the selective ligninolytic fungus ceriporiopsis subvermispora during growth on aspen wood reveal this organism's strategy for degrading lignocellulose. Appl Environ Microbiol. 2014;80:2062–70.
38. Tello M, Vergara F, Spencer E. Genomic adaptation of the ISA virus to Salmo salar codon usage. Virol J. 2013;10:223. doi:10.1186/1743-422X-10-223.
39. Aragonès L, Guix S, Ribes E, Bosch A, Pintó RM. Fine-tuning translation kinetics selection as the driving force of codon usage bias in the hepatitis A virus capsid. PLoS Pathog. 2010;6:e1000797. doi:10.1371/journal.ppat.1000797.
40. Plotkin JB, Kudla G. Synonymous but not the same: the causes and consequences of codon bias. Nat Rev Genet. 2011;12:32–42. doi:10.1038/nrg2899.
41. Sharp PM, Tuohy TM, Mosurski KR. Codon usage in yeast: cluster analysis clearly differentiates highly and lowly expressed genes. Nucleic Acids Res. 1986;14:5125–43.
42. Sharp PM, Cowe E, Higgins DG, Shields DC, Wolfe KH, Wright F. Codon usage patterns in Escherichia coli, Bacillus subtilis, Saccharomyces cerevisiae, Schizosaccharomyces pombe, Drosophila melanogaster and Homo sapiens; a review of the considerable within-species diversity. Nucleic Acids Res. 1988;16:8207–11.
43. Bahir I, Fromer M, Prat Y, Linial M. Viral adaptation to host: a proteome-based analysis of codon usage and amino acid preferences. Mol Syst Biol. 2009;5:311. doi:10.1038/msb.2009.71.
44. Angov E. Codon usage: nature's roadmap to expression and folding of proteins. Biotechnol J. 2011;6:650–9. doi:10.1002/biot.201000332.
45. Mueller S, Coleman JR, Papamichail D, Ward CB, Nimnual A, Futcher B, et al. Live attenuated influenza virus vaccines by computer-aided rational design. Nat Biotechnol. 2010;28:723–6. doi:10.1038/nbt.1636.
46. Zhou M, Guo J, Cha J, Chae M, Chen S, Barral JM, et al. Non-optimal codon usage affects expression, structure and function of clock protein FRQ. Nature. 2013;495:111–5. doi:10.1038/nature11833.
47. Novoa EM, Pavon-Eternod M, Pan T, Ribas de Pouplana L. A role for tRNA modifications in genome structure and codon usage. Cell. 2012;149:202–13. doi:10.1016/j.cell.2012.01.050.
48. Wenke T, Döbel T, Sörensen TR, Junghans H, Weisshaar B, Schmidt T. Targeted identification of short interspersed nuclear element families shows their widespread existence and extreme heterogeneity in plant genomes. Plant Cell. 2011;23:3117–28. doi:10.1105/tpc.111.088682.
49. Enav H, Béjà O, Mandel-Gutfreund Y. Cyanophage tRNAs may have a role in cross-infectivity of oceanic Prochlorococcus and Synechococcus hosts. ISME J. 2012;6:619–28. doi:10.1038/ismej.2011.146.
50. van Weringh A, Ragonnet-Cronin M, Pranckeviciene E, Pavon-Eternod M, Kleiman L, Xia X. HIV-1 modulates the tRNA pool to improve translation efficiency. Mol Biol Evol. 2011;28:1827–34. doi:10.1093/molbev/msr005.
51. Deswal D, Khasa YP, Kuhad RC. Optimization of cellulase production by a brown rot fungus Fomitopsis sp. RCK2010 under solid state fermentation. Bioresour Technol. 2011;102:6065–72. doi:10.1016/j.biortech.2011.03.032.
52. Levin L, Melignani E, Ramos AM. Effect of nitrogen sources and vitamins on ligninolytic enzyme production by some white-rot fungi. Dye decolorization by selected culture filtrates. Bioresour Technol. 2010;101:4554–63. doi:10.1016/j.biortech.2010.01.102.

53. Heinemann IU, Jahn M, Jahn D. The biochemistry of heme biosynthesis. Arch Biochem Biophys. 2008;474:238–51. doi:10.1016/j.abb.2008.02.015.
54. Ritch TG, Gold MH. Characterization of a highly expressed lignin peroxidase-encoding gene from the basidiomycete Phanerochaete chrysosporium. Gene. 1992;118:73–80.
55. Floudas D, Binder M, Riley R, Barry K, Blanchette RA, Henrissat B, et al. The Paleozoic origin of enzymatic lignin decomposition reconstructed from 31 fungal genomes. Science. 2012;336:1715–9. doi:10.1126/science.1221748.
56. **- Tables**.

Tables

Table 1
Statistics of tRNAs present in the C. subvermispora genome

tRNA Type	Number of tRNAs	Anticodon and Frecuency	Scaffold	Number of tRNAs with Introns
Ala	14	AGC[7], CGC[5], TGC[2]	3,5,7,12,14,18,19	13
Arg	16	ACG[8], CCG[2], CCT[2], TCG[3], TCT[1]	2,3,5,9,14,19,31,39	15
Asn	5	GTT[5]	4,5,7,13,14	5
Asp	10	GTC[10]	1,6,8,12,16,19,28,41	7
Cys	3	GCA[3]	10,13,20	1
Gln	8	CTG[6], TTG[2]	1,2,6,17,18,28	8
Glu	15	CTC[12], TTC[3]	3,8,9,19,28	5
Gly	17	CCC[2], GCC[12], TCC[3]	1,2,5,7,9,10,13,20	1
His	5	GTG[5]	2,6,7,12	5
Ile	9	AAT[8], TAT[1]	4,5,21,33	9
Leu	15	AAG[7], CAA[2], CAG[4], TAA[1], TAG[1]	1,2,4,5,6,7,9,13,17,25,27,30	14
Lys	11	CTT[10], TTT[1]	1,3	11
Met	8	CAT[8]	1,2,3,9,10,11,12,23	7
Phe	5	GAA[5]	1,5,9,10	4
Pro	10	AGG[5], CGG[4], TGG[1]	4,6,7,14,18,22,35	10
Ser	13	ACT[1], AGA[4], CGA[3], GCT[3], TGA[2]	1,3,4,5,6,9,15,18	8
Thr	9	AGT[6],CGT[2], TGT[1]	1,3,4,8,9,10,22	6
Trp	3	CCA[3]	19,20,30	3
Tyr	4	GTA[4]	2,3,8,30	4
Val	11	TAC[1], CAC[2], AAC[6], CAC[1], AAC[1]	5,6,12,14,15,28,32	2

A12

Table 2

Codon frecuency in the C. subvermispora genome and availability of their cognate tRNAs.

#Codon	AA	Fraction	Frequency	Number	RSCU	W (tRNA)	w (tRNA)
GCA	A	0.238	21.928	113366	0.950	5.500	0.407
GCC	A	0.237	21.845	112935	0.950	3.500	0.259
GCG	A	0.323	29.736	153734	1.290	6.000	0.444
GCT	A	0.202	18.609	96206	0.810	7.000	0.519
TGC	C	0.648	8.524	44068	1.300	3.000	0.222
TGT	C	0.352	4.631	23941	0.700	1.500	0.111
GAC	D	0.597	33.526	173324	1.190	10.000	0.741
GAT	D	0.403	22.634	117017	0.810	5.000	0.370
GAA	E	0.343	20.056	103689	0.690	3.000	0.222
GAG	E	0.657	38.500	199038	1.310	13.500	1.000
TTC	F	0.739	26.294	135936	1.480	5.000	0.370
TTT	F	0.261	9.298	48072	0.520	2.500	0.185
GGA	G	0.205	13.330	68916	0.820	3.000	0.222
GGC	G	0.390	25.415	131394	1.560	12.000	0.889
GGG	G	0.186	12.115	62634	0.740	3.500	0.259
GGT	G	0.219	14.291	73883	0.880	6.000	0.444
CAC	H	0.575	14.936	77216	1.150	5.000	0.370
CAT	H	0.425	11.041	57079	0.850	2.500	0.185
ATA	I	0.158	7.411	38314	0.470	5.000	0.370
ATC	I	0.562	26.331	136128	1.690	4.000	0.296
ATT	I	0.279	13.086	67652	0.840	8.000	0.593
AAA	K	0.270	10.918	56446	0.540	1.000	0.074
AAG	K	0.730	29.453	152267	1.460	10.500	0.778
CTA	L	0.088	8.184	42311	0.530	4.500	0.333
CTC	L	0.329	30.487	157614	1.980	3.500	0.259
CTG	L	0.240	22.203	114786	1.440	4.500	0.333
CTT	L	0.155	14.379	74337	0.930	7.000	0.519
TTA	L	0.045	4.168	21550	0.270	1.000	0.074
TTG	L	0.142	13.148	67976	0.850	2.500	0.185
ATG	M	1.000	21.228	109746	1.000	8.000	0.593
AAC	N	0.623	19.336	99964	1.250	5.000	0.370
AAT	N	0.377	11.702	60500	0.750	2.500	0.185
CCA	P	0.210	13.766	71171	0.840	3.500	0.259
CCC	P	0.259	16.990	87835	1.030	2.500	0.185
CCG	P	0.296	19.453	100572	1.180	4.500	0.333
CCT	P	0.235	15.465	79953	0.940	5.000	0.370
CAA	Q	0.365	13.696	70805	0.730	2.000	0.148
CAG	Q	0.635	23.829	123194	1.270	7.000	0.519
AGA	R	0.096	6.328	32716	0.580	1.000	0.074
AGG	R	0.116	7.632	39455	0.690	2.500	0.185
CGA	R	0.137	9.055	46811	0.820	7.000	0.519

#Codon	AA	Fraction	Frequency	Number	RSCU	W (tRNA)	w (tRNA)
CGC	R	0.312	20.574	106367	1.870	4.000	0.296
CGG	R	0.172	11.345	58653	1.030	3.500	0.259
CGT	R	0.168	11.059	57173	1.010	8.000	0.593
AGC	S	0.187	15.803	81699	1.120	3.000	0.222
AGT	S	0.093	7.846	40561	0.560	2.500	0.185
TCA	S	0.132	11.104	57406	0.790	4.000	0.296
TCC	S	0.197	16.574	85687	1.180	2.000	0.148
TCG	S	0.225	18.939	97912	1.350	4.000	0.296
TCT	S	0.167	14.052	72645	1.000	4.000	0.296
TAA	Stop	0.215	0.472	2440	0.640	NA	NA
TAG	Stop	0.301	0.661	3416	0.900	NA	NA
TGA	Stop	0.484	1.063	5493	1.450	NA	NA
ACA	T	0.229	13.700	70828	0.920	4.000	0.296
ACC	T	0.270	16.133	83408	1.080	3.000	0.222
ACG	T	0.303	18.161	93891	1.210	2.500	0.185
ACT	T	0.198	11.852	61271	0.790	6.000	0.444
GTA	V	0.127	8.265	42728	0.510	4.500	0.333
GTC	V	0.399	25.883	133812	1.590	3.500	0.259
GTG	V	0.289	18.780	97088	1.160	3.500	0.259
GTT	V	0.185	12.013	62104	0.740	7.000	0.519
TGG	W	1.000	14.757	76294	1.000	3.000	0.222
TAC	Y	0.627	16.316	84352	1.250	4.000	0.296
TAT	Y	0.373	9.691	50102	0.750	2.000	0.148

Table 3

Correlation coefficient between gene expression, codon bias and translational efficiency.

		BMA (ρ)	Glu(ρ)	BMA/Glu (ρ)	BMA/Glu(ρ)[a]	BMA/Glu(ρ)[b]	BMA/Glu(ρ)[c]
	CAI	-4.95×10^{-3}	$-8.04 \times 10^{-2***}$	$3.39 \times 10^{-1***}$	$4.20 \times 10^{-1**}$	$4.66 \times 10^{-1***}$	$3.09 \times 10^{-1***}$
	CPB	$6.09 \times 10^{-2***}$	$6.59 \times 10^{-3***}$	$2.61 \times 10^{-1***}$	NA	$3.26 \times 10^{-1***}$	$2.47 \times 10^{-1***}$
	Nc	$1.81 \times 10^{-2*}$	$8.17 \times 10^{-2***}$	$-2.63 \times 10^{-1***}$	$-4.51 \times 10^{-1***}$	$-4.42 \times 10^{-1***}$	$-2.21 \times 10^{-1***}$
	tAI	$3.48 \times 10^{-2***}$	$-3.52 \times 10^{-2***}$	$3.17 \times 10^{-1***}$	$2.84 \times 10^{-1*}$	$4.36 \times 10^{-1***}$	$2.82 \times 10^{-1***}$
	AAtAI	$6.54 \times 10^{-2***}$	8.34×10^{-4}	$2.67 \times 10^{-1***}$	$4.75 \times 10^{-1***}$	$4.17 \times 10^{-1***}$	$2.27 \times 10^{-1***}$
*	p < 0.05						
**	p < 0.01						
***	p < 0.001						
a	Genes with p values lower than 0.001, n = 56						
b	Genes with p values between 0.001 and 0.05, n = 1572						
c	Genes with p values higher than 0.05, n = 10471						
Glu	Glucose						

Table 4

Adaptation to codon usage, tRNA pool and codon bias of genes from Ceriporiopsis subvermispora involved in mineralization and digestion of lignocellulose. I highest values

Ligninolytic Function	ID	CAI	AAtAI	tAI	CPB	Nc	Z-CAI	Z-AAtAI	Z-CPB	Z-tAI	Putative function	Microarray[4] signal (log2) Cesubvub Glucose	Microarray[4] signal (log2) Cesubvub BMA
Fungal lignin peroxidase	49863	0,814	0,720	0,371	0,028	40,95	1,236	1,848	0,136	1,513	Peroxidase	9,58	11,15
	50297	0,822	0,735	0,384	0,055	41,46	1,394	2,321	1,565	2,111	Peroxidase	13,13	11,88
	50686	0,823	0,728	0,379	0,051	41,50	1,414	2,100	1,387	1,901	Peroxidase	9,35	9,42
	106380	0,815	0,716	0,371	0,039	38,67	1,255	1,722	0,705	1,536	Peroxidase	8,93	8,93
	111364	0,770	0,689	0,357	0,015	50,57	0,364	0,872	-0,611	0,929	Peroxidase, VP-like	9,18	9,11
	117521	0,819	0,722	0,367	0,042	40,88	1,335	1,911	0,876	1,340	Peroxidase	9,57	9,49
	124144	0,696	0,653	0,321	0,031	61,00	-1,101	-0,261	0,284	-0,660	Peroxidase	11,30	11,21
	126018	0,828	0,721	0,366	0,030	40,55	1,513	1,880	0,230	1,314	Peroxidase	9,69	9,50
	126023	0,814	0,717	0,352	0,020	39,56	1,236	1,754	-0,311	0,690	Peroxidase	9,79	9,55
	126058	0,788	0,673	0,344	0,026	44,68	0,721	0,368	-0,015	0,340	Peroxidase	9,94	9,44
	128590	0,824	0,724	0,367	0,052	41,12	1,434	1,974	1,409	1,352	Peroxidase	9,78	9,52
	129418	0,767	0,709	0,373	0,044	49,06	0,305	1,502	1,011	1,619	Peroxidase	10,22	12,92
	130659	0,781	0,721	0,368	0,022	50,53	0,582	1,880	-0,220	1,396	Peroxidase, LiP-like	10,88	10,30
	136058	0,770	0,688	0,352	0,029	50,91	0,364	0,841	0,167	0,683	Peroxidase	10,16	10,76
	151947	0,831	0,729	0,373	0,052	40,42	1,572	2,132	1,410	1,600	Peroxidase	9,00	8,93
	155372	0,797	0,711	0,369	0,040	42,88	0,899	1,565	0,765	1,443	Peroxidase	8,86	8,84
	169968	0,816	0,718	0,369	0,032	42,96	1,275	1,785	0,359	1,434	Peroxidase	10,27	10,16
Laccase	84170	0,741	0,664	0,325	0,032	58,64	-0,210	0,085	0,338	-0,517	laccase	10,45	10,34
	88089	0,841	0,684	0,336	0,039	41,00	1,770	0,715	0,695	-0,016	Laccase	9,38	9,29
	120834	0,737	0,669	0,330	0,011	57,80	-0,289	0,242	-0,825	-0,271	Laccase	10,74	10,66
	127045	0,752	0,681	0,339	0,024	54,32	0,008	0,620	-0,103	0,112	Laccase	10,54	10,02
	127050	0,721	0,667	0,334	0,015	60,26	-0,606	0,179	-0,581	-0,098	Laccase	11,35	11,19
	130783	0,791	0,705	0,347	0,037	43,13	0,780	1,376	0,605	0,461	Laccase	11,02	13,77
	149668	0,775	0,679	0,334	0,019	51,25	0,463	0,557	-0,397	-0,120	Laccase	9,91	9,81
Ligninolytic Function	**ID**	**CAI**	**AAtAI**	**tAI**	**CPB**	**Nc**	**Z-CAI**	**Z-AAtAI**	**Z-CPB**	**Z-tAI**	**Putative function**	**Microarray[4] signal (log2) Cesubvub Glucose**	**Microarray[4] signal (log2) Cesubvub BMA**
Cellulose Binding Protein	59733	0,812	0,704	0,346	0,026	46,11	1,196	1,344	-0,018	0,414	GH10	9,24	13,69
	66688	0,824	0,716	0,347	0,025	47,02	1,434	1,722	-0,048	0,488	GH61-CBM1	9,62	14,87
	67561	0,853	0,730	0,366	0,020	39,43	2,008	2,163	-0,336	1,313	GH10	10,30	13,71
	68569	0,837	0,735	0,345	0,024	43,59	1,691	2,321	-0,095	0,383	CE1	9,80	13,10
	79557	0,802	0,715	0,345	0,022	42,15	0,998	1,691	-0,231	0,373	GH5	10,24	14,04
	87580	0,783	0,693	0,321	0,010	49,31	0,622	0,998	-0,884	-0,692	CE16	10,94	14,46
	89533	0,838	0,736	0,353	0,052	42,06	1,711	2,352	1,426	0,719	GH61-CBM1	10,35	13,56

Ligninolytic Function	ID	CAI	AAtAI	tAI	CPB	Nc	Z-CAI	Z-AAtAI	Z-CPB	Z-tAI	Putative function	Microarray[4] signal (log2) Cesubvub Glucose	Microarray[4] signal (log2) Cesubvub BMA
	89534	0,870	0,731	0,352	0,044	37,58	2,345	2,195	0,980	0,693	GH61-CBM1	9,42	10,13
	101925	0,845	0,723	0,359	0,033	38,22	1,850	1,943	0,394	1,020	GH7-CBM1	8,84	8,96
	106777	0,804	0,712	0,348	0,035	49,87	1,038	1,596	0,511	0,500	GH5	9,54	14,35
	109840	0,857	0,741	0,370	0,039	38,16	2,087	2,509	0,733	1,471	GH10	9,52	11,85
	129028	0,852	0,733	0,363	0,042	39,91	1,988	2,258	0,849	1,181	GH5	9,81	10,79
	133809	0,865	0,713	0,338	0,040	37,08	2,246	1,628	0,779	0,091	GH11-CBM1	10,86	11,42
	148588	0,856	0,742	0,376	0,008	34,45	2,067	2,541	-0,967	1,756	GH7-CBM1	11,02	12,62
CDH	84792	0,803	0,688	0,338	0,023	48,09	1,018	0,841	-0,137	0,071	CDH	9,29	13,76
	87110	0,769	0,679	0,332	0,024	50,71	0,345	0,557	-0,111	-0,189		11,31	11,04
	125610	0,762	0,665	0,318	0,044	53,34	0,206	0,116	0,981	-0,814	cir1 CBM1	10,17	10,40
	147544	0,712	0,665	0,318	0,025	56,98	-0,784	0,116	-0,060	-0,798		11,24	11,02
Delta 12 Dehidrogenase	58880	0,727	0,670	0,317	0,008	59,40	-0,487	0,274	-0,973	-0,853	Delta 12 Fatty acid desaturase	10,36	10,29
	121074	0,731	0,644	0,313	0,022	55,49	-0,408	-0,545	-0,219	-1,033	Delta 12 Fatty acid desaturase	10,58	10,23
	124050	0,856	0,736	0,368	0,042	38,22	2,067	2,352	0,875	1,416	Delta-12 fatty acid desaturase, Cs-fad2	12,74	12,66
	136101	0,714	0,649	0,310	0,007	56,78	-0,745	-0,387	-1,016	-1,178	Delta 12 Fatty acid desaturase	11,01	12,54
	167690	0,736	0,670	0,325	0,013	58,45	-0,309	0,274	-0,706	-0,484	Delta 12 Fatty acid desaturase	10,67	10,11
Delta 9 Dehidrogenase	87875	0,810	0,694	0,343	0,035	44,18	1,156	1,030	0,475	0,313	Delta-9 fatty acid desaturase, Cs-ole1	8,93	8,94
	129045	0,728	0,679	0,335	0,045	56,92	-0,467	0,557	1,017	-0,075	Delta-9 fatty acid desaturase, Cs-ole1	8,95	8,91
	129048	0,848	0,740	0,366	0,050	41,46	1,909	2,478	1,291	1,292	Delta-9 fatty acid desaturase, Cs-ole1	11,78	12,35
	133675	0,760	0,669	0,330	0,024	54,62	0,166	0,242	-0,113	-0,266	Delta-9 fatty acid desaturase, Cs-ole1	9,64	9,51

Figures

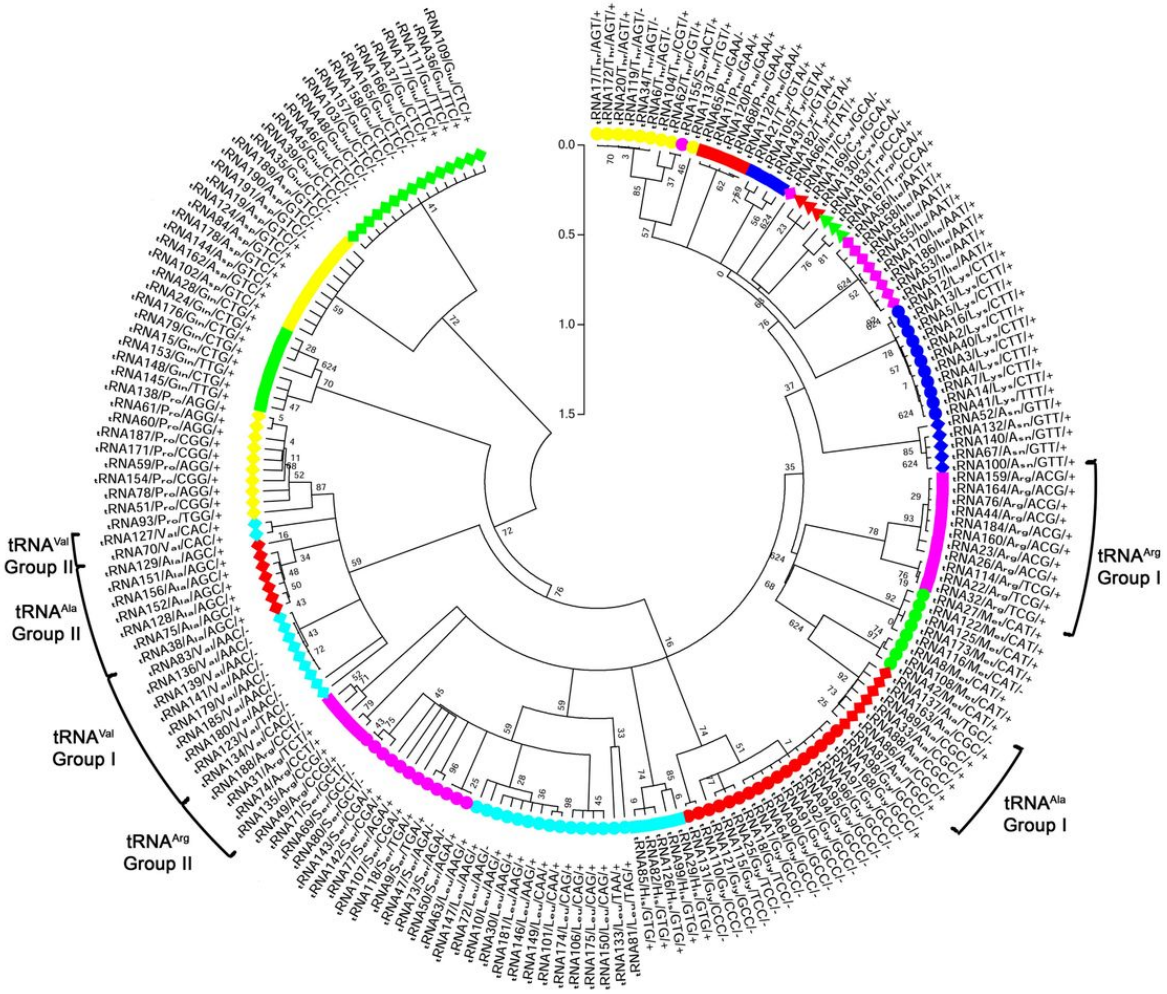

Figure 1

Phylogenetic analysis of tRNAs from C. subvermispora. tRNAs were numbered according the tRNAScan-SE output. Prediction of aminoacid charge, sequence of anticodon and presence (+) or absence (-) of introns are indicated. Each tRNA type is indicated with a different symbol and color.

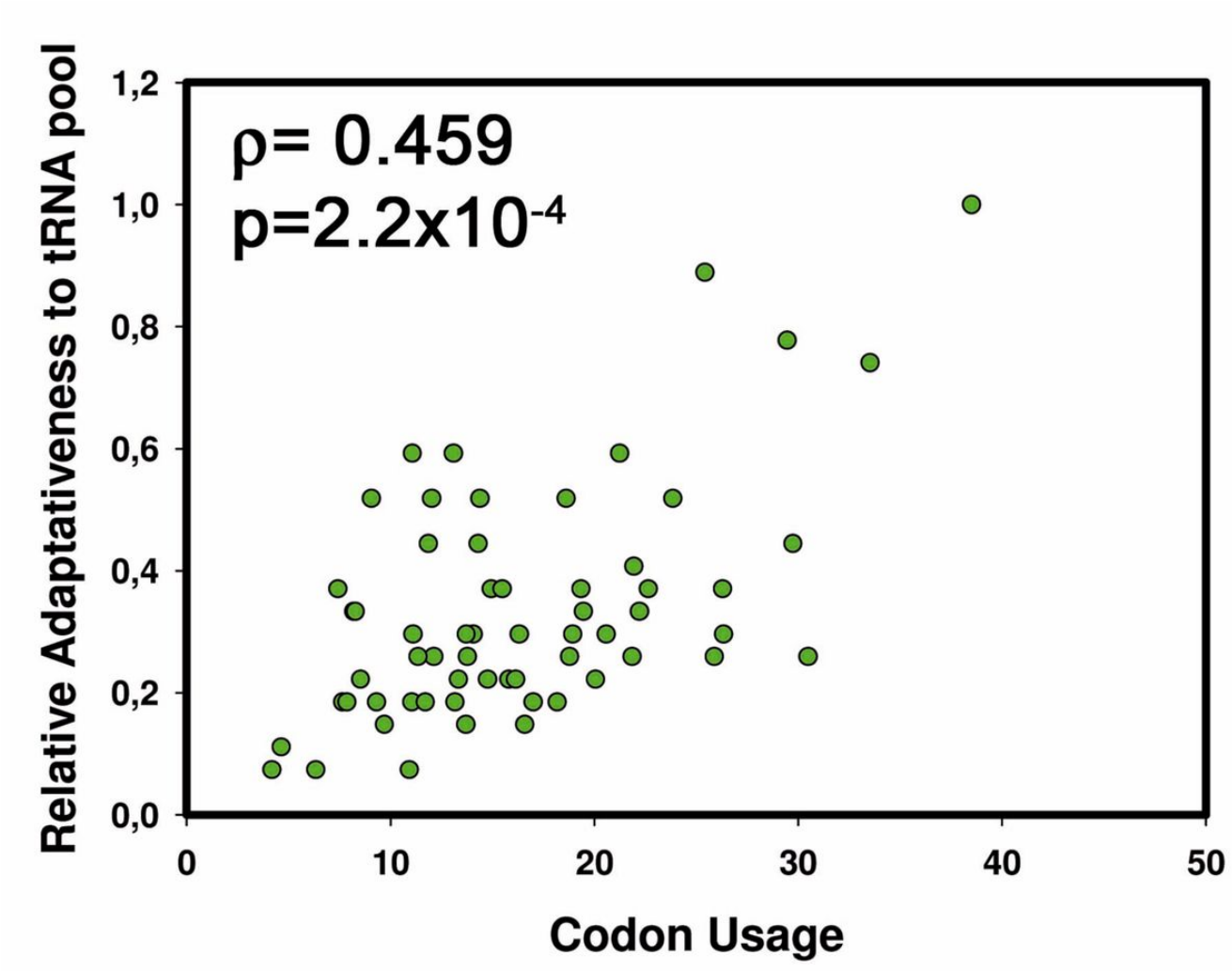

Figure 2

Correlation between codon usage and relative adaptativeness to the tRNA pool in C. subvermispora. Correlation between the relative adaptativeness to the tRNA pool and codon frequency usage (per thousand) of the 61 codons is shown.

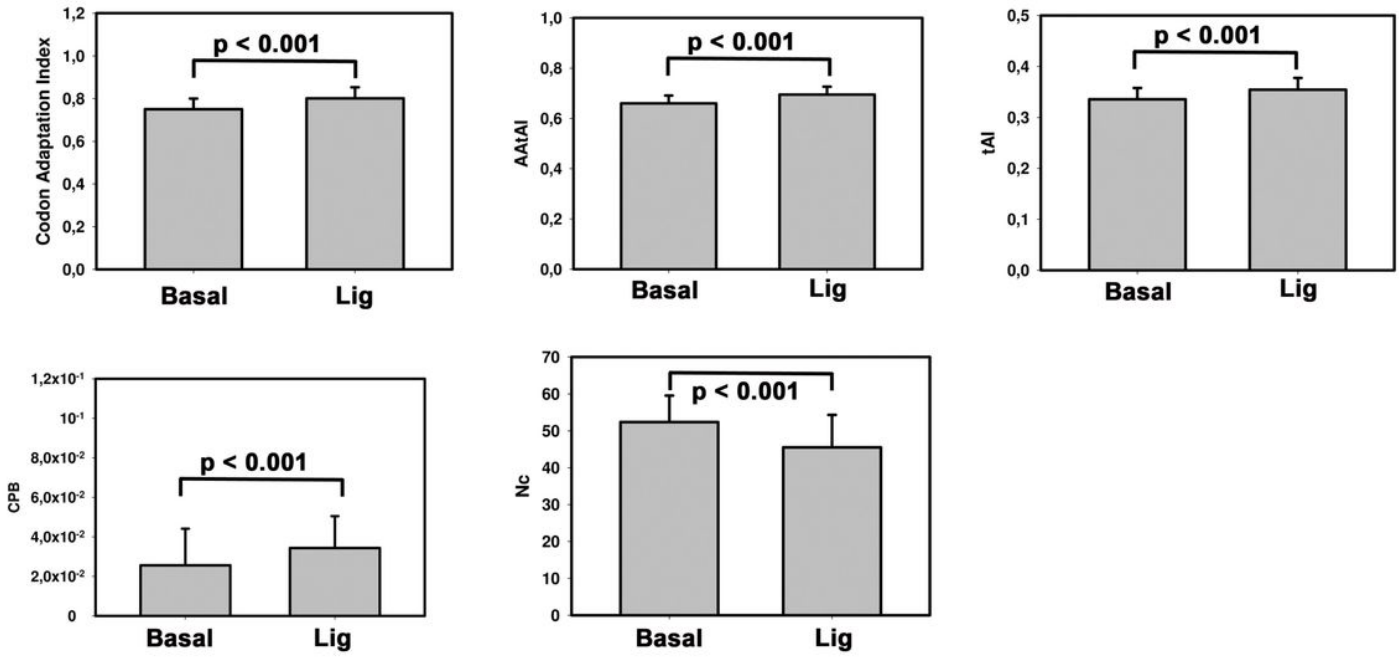

Figure 3

Codon bias and adaptation to tRNA pool of ligninolytic and non-ligninolytic genes. The mean values of CAI, CPB, AAtAI, tAI, and Nc of genes that do not modify their expression in response to BMA (Basal) and of those genes whose expression is induced at least twice (Lig) (Group B) are shown. Statistical differences between both groups were assessed by non-parametric test using a p value lower than 0.001.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- Additionalfile5.xls
- Additionalfile4.xls
- Additionalfile3.xls
- Additionalfile2.xls
- Additionalfile1.xls