A full complement of transfer RNA genes in the *Candida albicans* genome

J. Fan[*,1], **F. Savard**[1], **M. Wu**[1], **and S.-H. Shen**[*,1,2]

[1]Mammalian Cell Genetics Group, Health Sector, Biotechnology Research Institute, National Research Council of Canada, Montreal, Quebec H4P 2R2, Canada
[2]Department of Medicine, McGill University, Montreal, Quebec H3G 1A4, Canada

The functions of non-coding genes, such as transfer RNA (tRNA) genes, in *Candida albicans* have been largely overlooked, though it is widely realized that they play important roles in regulating gene expression. As an initial effort to study small RNA genes from *C. albicans*, we have identified a total of 132 tRNA genes from the nuclear genome using tRNAscan-SE. These tRNA genes represent 43 unique tRNA gene species decoding all of the 20 amino acids. Together with the previously annotated 30 tRNA genes, of which 24 are unique, from the mitochondrial genome of *C. albicans*, a total of 162 tRNA genes were used in this study. Since the "universal" leucine codon CTG specially encodes for serine in *C. albicans*, this non-standard translational event may add some special features to the tRNA genes. Phylogenetic relationships of tRNA genes from the nuclear and mitochondrial genomes were compared to demonstrate some recent gene duplications and interorganellar gene transfer. The correlation of codon usage with copy numbers of the tRNA genes was also discussed.

Key words: tRNA genes; phylogenetic analysis; codon usage; *Candida albicans*
Abbreviations: tRNAs, transfer RNAs; ORF, open reading frame, HGT, horizontal gene transfer
Running title: tRNA genes in *Candida albicans*

1. Introduction

Small non-coding RNAs have recently drawn broad attention to their potential functions in animals and plants [1,2]. It was reported that 98% of all transcriptional outputs in humans are non-coding RNAs [3]. Especially, small interference RNAs (siRNAs) were found to function in silencing cellular gene expression *in vivo* and have been genome-widely used in the study of gene functions [4]. Further, even untranslated RNAs (introns) were found to cause human diseases [5]. Since the majority of the eukaryotic genomes comprise intronic regions, repeat sequences, and other assorted "junk DNAs", new research on non-coding RNAs will shed more light on the dark side of the genome than its bright side, the encoding genes, which can be easily annotated through comparative genomics. Among the non-coding RNAs, there are tRNA-like RNA, mRNA-like RNA (tmRNA), microRNA (miRNA), small interference RNA (siRNA), small nucleolar RNA (snoRNA), small nuclear RNA (snRNA), ribosomal RNA and transfer RNA (tRNA). Since most of their genomic sequences or related maintenance/regulation mechanisms were evolutionarily conservative, it allows us to exploit their functions based on the genomic data collected from previous studies.

Candida albicans, a human commensal pathogenic fungus, is dimorphic with a unicellular budding yeast form and a filamentous form, which makes it an ideal model organism for studying other pathogenic fungi, and the interactions between those pathogens and their human host [6]. Since the genome sequence of *Candida albicans* is nearly completed (http://www-sequence.stanford.edu/group/candida/ index.html), and valuable bioinformatic tools are available [7], it is now possible to survey the putative tRNAs of *C. albicans* in genome-wide. The complete collection of tRNAs reported previously from the model single-celled eukaryote *Saccharomyces cerevisiae* has illustrated the minimal complement of tRNA genes, the relationship between tRNA species copy

[*] Corresponding author: email: jinjiang.fan@gmail.com or Shi.Shen@nrc.ca, Phone: 1-514-496 6318, Fax: 1-514-496 6319

number, intracellular tRNA concentration and protein codon usage for an eukaryote [8,9]. Based on the modified wobble rules proposed for *S. cerevisiae* by Guthrie & Abelson [10], 46 different tRNA species are predicted in yeast. The reported set of 43 tRNA families for the yeast is very close to this prediction [8]. A similar scenario could be seen for the human pathogenic yeast *C. albicans*. However, in *Candida spp.* the standard leucine-CTG codon changes to a serine [11-13]. This unique genetic code reassignment suggests that tRNA genes in *C. albicans* have alternative species-specific roles in protein biosynthesis through the direct interaction of a tRNA with its amino acid in the synthesized protein. To further understand the mechanism of translation system and beyond, it is important to comprehend the whole set of tRNAs for this human pathogenic yeast.

As an initial effort to study tRNA genes in *C. albicans*, here we report 132 tRNA genes from the *C. albicans* nuclear genomes and study their phylogenetic relationship with the tRNA genes from mitochondria, and the correlation of their copy numbers with codon usages.

2. Materials and Methods

2.1 *Sequences and database searches*. The *C. albicans* tRNA genes were identified from the *C. albicans* genome database assembly (Version 19) at the Stanford DNA Sequencing and Technology Center (http://www-sequence.stanford.edu/group/candida) using tRNAscan-SE program with eukaryote-specific parameters [7]. The number of the tRNA genes was estimated based upon occurrence within the 10.4X sequence coverage in contigs. The mitochondrial tRNA sequences were extracted according to the annotation of the whole mitochondria sequence of *C. albicans* (AF285361).

2.2 *Multiple sequence alignment and codon usage estimation*. The tRNA sequences from nucleus and mitochondria were aligned using the Clustal W program [14]. After initial alignments, all tRNAs sequences were then manually optimized to force the anticodon loop aligned. Encoding sequences from 9259 ORFs and all contigs were totally analyzed for the estimation of codon usage using program CUSP in EMBOS (http://www.hgmp.mrc.ac.uk/Software/EMBOSS). Intron sequences from corresponding predicted tRNAs were manually adjusted to show their identities.

2.3 *Phylogenetic tree construction*. After removing all sites containing alignment gaps and missing-information, the conserved regions of all aligned tRNAs from both the nuclear and mitochondrial genomes of *C. albicans* were used to generate unrooted maximum-likely trees with majority rule and bootstrap analysis by using program SEQBOOT, DNAML and CONSENCE in the PHYLIP package [15]. The neighbor-joining tree was constructed in MEGA 2.1 [16] using Jukes-Cantor model. The statistical significance of the phylogenetic tree obtained was tested by bootstrap analysis with 100 replicates of random addition using the SEQBOOT program in PHYLIP. The resulting trees were visualized with the program TreeView [17] and with the program MEGA [16].

3. Results and Discussion

3.1 *Genomic survey of the tRNA genes from* C. albicans. It is well know that tRNAs play a critical role in protein synthesis as they act as adaptor molecules in fidelity to transfer the genetic information from messenger RNA (mRNA) to protein sequence. To identify the whole set of the tRNA genes from the *C. albicans* haploid genome, an analysis based on the computer-aided program tRNAscan-SE was performed [7], resulting in completely identification of all genes encoding tRNA species. A total of 132 nuclear-encoded tRNA genes were detected using this program (Fig. 1), and 43 unique tRNA gene species were identified and aligned to show their identities (Fig. 2a). Our prediction is very closely to the total number of tRNA genes described in the CandidaDB (http://genolist.pasteur.fr/CandidaDB) [18]. The triplet anticodons were manually adjusted, and were marked in a blank box. Sixty-nine out of 132 tRNA genes possess an intron located right after the first nucleotide following each triplet anticodon (Table 1). Since some features of tRNA genes in mitochondria, such as general cloverleaf structure, are by far less well known and may deviate from nuclear-encoding tRNAs, these sequences were directly

extracted from the annotated *C. albicans* mitochondrial genome (AF285261/NC_002653). Twenty-four unique species out of 30 tRNA genes from the mitochondrial genome were presented in an alignment (Fig. 2b). The highly conserved sequences of the two potential polymerase III-binding sites in these tRNA genes are aligned and schematically shown as Box A and Box B.

3.2 *The redundancy of tRNA families and their tRNA genes.* Among 132 nuclear tRNA genes, 43 possess unique tRNA sequences characteristic of distinct codon specificities with exception of $tRNA^{Met}$ (Table 1). Based on "wobble hypothesis" and the modified wobble rules proposed from the observed yeast tRNA species [10,19], the first position of the tRNA anticodon could pair with more than one possible nucleotide in the third position of mRNA codons. In brief, an anticodon G at the first position could pair with U or C, a U with A or G, and an I (inosine, which is post- transcriptionally modified through hydrolytic deamination from the adenosines in the first position of tRNA anticodons) with U, C, or A [20]. Accordingly, the 43 tRNA gene families in *C. albicans* well meet requirements to translate all the possible codons (Fig. 4). Two $tRNA^{Met}$ genes have been proposed to have initiation or elongation function during protein synthesis [9]. An intron with more or less sequence variation was found in 69 out of 132 tRNA genes, and intron sequence deviations were also observed in some tRNAs that have identical spliced sequences, suggesting that these members of the tRNA genes were generated through very recent gene duplication. Some intron sequences were served as part of the encoding genes as shown in Fig 3a. Interestingly, while one copy of the $tRNA^{Lys}$ gene possesses one small intron seemly composed of a repeated sequence, the rest four copies of this tRNA gene have no intron. In general, a remarkable sequence similarity was observed within each gene families, whereas introns maintained in tRNA genes are diverse among different tRNA families. Since GC contents of these introns were 32.2%, similar to the average GC content (32.9%) of the genome, it is difficult to determine whether they possess any mobile genetic elements with a lower GC content as proposed previously [21]. Clearly, intron-retention within each tRNA gene family indicates that positive selection related to functions may have promoted an indefinite intron preservation as shown above [22,23]. However, most tRNA gene species maintained a variable intron sequence though some completely lost intron, such as $tRNA^{Lys}$ genes (Fig. 3a). This raises the question as to whether it is necessary to maintain an intron within a specific tRNA gene for correct addition of anticodon base modifications. Nevertheless, the variation of intron sequences within the same tRNA gene families may indicate the history of tRNA gene redundancy through the role of introns within tRNAs.

Table 1. Gene copy number and predicted decoding specificities of the tRNA genes from *C. albicans*

tRNA species	tDNA anticodon	Number of tRNA genes	tRNA[a] anticodon	Codon[b] preference	Intron (Y/N)	Accession No.[c]	Reference
Ala	AGC	7	IGC	GCU, GCC	N	CA7044	[24]
Ala	TGC	2	UGC	GCA, GCG	N	CA7063	
Arg	CCG	1	CCG	CGG, CGU, CGA, CGG	Y	CA7069	
Arg	CCT	1	CCU	AGG	Y	CA7033	
Arg	TCT	5	UCU	AGA	N	CA7004	
Arg	ACG	2	ICG	CGU,CGC,CGA	Y	CA7120	
Asn	GTT	4	GUU	AAC, AAU	Y	CA7009	
Asp	GTC	7	GUC	GAU, GAC	N	CA7007	[24]
Cys	GCA	2	GCA	UGU, UGC	Y	CA7093	
Gln	CTG	1	CUG	CAG	N	CA7038	
Gln	TTG	5	UUG	CAA	Y	CA7026	
Glu	CTC	1	CUC	GAG	N	CA7019	
Glu	TTC	7	UUC	GAA	Y	CA7006	

Gly	CCC	1	CCC	GGG	N	CA7078	
Gly	GCC	6	GCC	GGC, GGU	N	CA7013	
Gly	TCC	2	NCC	GGA	N	CA7049	
His	GTG	2	GUG	CAU, CAC	Y	CA7014	
Ile	AAT	5	IAU	AUU, AUC	Y	CA7001	[24]
Ile	TAT	1	UAU	AUA	Y	CA7047	
Leu	AAG	2	AAG	CUU, CUC CUA	Y	CA7051	
Leu	CAA	6	CAA	UUG	Y	CA7027	
Leu	TAA	5	NAA	UUA	Y	CA7041	
Lys	CTT	2	CUU	AAG	Y	CA7005	
Lys	TTT	4	UUU	AAA	N	CA7031	
Lys	TTT	1	UUU	AAA	Y	CA7101	
Met	CAT	2	CAU	AUG	N	CA7002	
Met	CAT	2	CAU	AUG	N	CA7077	
Phe	GAA	5	GAA	UUC, UUU	Y	CA7003	
Pro	AGG	1	IGG	CCU, CCC	Y	CA7100	
Pro	TGG	5	UGG	CCA, CCG	Y	CA7011	
Ser	AGA	4	IGA	UCU, UCC	N	CA7012	
Ser*	CAG	1	CAG	CUG	N	CA7020	[25] [24]
Ser	CGA	1	CGA	UCG	Y	CA7088	
Ser	GCT	2	GCU	AGC, AGU	Y	CA7035	
Ser	TGA	3	UGA	UCA	N	CA7061	
Thr	AGT	6	IGU	ACU, ACC	N	CA7008	
Thr	TGT	2	UGU	ACA	N	CA7079	
Thr	CGT	1	CGU	ACG	Y	CA7117	
Trp	CCA	2	CCA	UGG	Y	CA7071	
Tyr	GTA	5	GUA	UAU, UAC	Y	CA7021	
Val	AAC	6	IAC	GUU, GUC	N	CA7028	
Val	CAC	1	CAC	GUG	N	CA7037	
Val	TAC	1	UAC	GUA	Y	CA7090	

a, tRNA anticodons were derived from tRNAscan-SE [7], except for the conversion of adenosine to inosine following the rule of Lim [26].
b, Codon preference was following the conventions proposed by Gurine and Abelson [10].
c. Each type of tRNA genes is reprehensively shown here as a database accession number in the CandidaDB (http://genolist.pasteur.fr/CandidaDB). Nucleotide sequence data for *Candida albicans* were obtained from the Stanford Genome Technology Center website at http://www-sequence.stanford.edu/group/candida. Sequencing of *C. albicans* was accomplished with the support of the NIDR and the Burroughs Wellcome Fund.
* Several publications related the ser-tRNA$_{CAG}$ are not cited because of space-limitation.

3.3 *Phylogeny of tRNAs from nucleus and mitochondria*. The CTG-anticodon of ser-tRNA in *C. albicans* was recently reported to be of original from a serine-tRNA and not from a leucine tRNA in the ancestor yeast species. This codon reassignment occurred approximately 170 million years ago [11]. However, the relationships of all tRNAs in *C. albicans* have not been genome-widely investigated yet. Based on 132 tRNA genes identified from the near complete genome sequences, 77.8% of the tRNA genes were dispersed in clusters from 2 to 7 copies in a contig (Fig. 1). We can see that the dispersal of most tRNA genes is nonrandom on a genome-wide scale because the numbers of the tRNA genes maintained in each contig were not correlated with its size, even though it is worthy of seeing a clear picture through mapping all the tRNA genes on chromosomes. Our phylogenetic analysis revealed some recent

evolutionary events of this molecule from *C. albicans*. Results clearly indicated that the tRNA genes were generated through gene duplication (Lys1/Lys2, Ala1/Ala2, Gln1/Gln2, Glu1/Glu2, Pro1/Pro2, Val1/Val2, and mtLeud/mtLeu2), intron procession (intron-containing tRNAs), interorganellar gene transfer between nucleus and mitochondria (mtRNAAsn and Thr1), and specific point mutation (SerL4) (Fig. 2c). These evolutionary events were strongly supported by bootstrap analysis. Gene duplication has been proposed as a necessary source of material for the origin of evolutionary novelties [27]. Here seven tRNA species (Ala, Gln, Glu, Lys, Pro, Ser, Val) from nucleus and one (Leu) from mitochondria seemed to have originated through a gene duplication following point mutation to increase their codon capacity. Interorganellar gene transfer between nucleus and mitochondria is very common in plants [28-30], but has not been reported so far in *C. albicans*. Through comparison, we found one case in *C. albicans*, *i.e.* interorganellar gene transfer between $tRNA^{Thr1}$ and mitochondrial $tRNA^{Asn}$, and more cases were also observed from other genes or sequences through a direct comparison of the near- and mitochondrial-genomes (data not shown) (Fig. 3b). To further characterize the relationship between $tRNA^{Thr1}$ and $tRNA^{Asn}$ from the nuclear- and mitochondrial- genomes, a Neighbor-joining tree was constructed based on the sequence alignment as shown in Fig 3b, and the result demonstrated that nuclear-encoding $tRNA^{Thr1}$ was of mitochondrial origin. In the case of horizontal gene transfer (HGT), the change from one isoaccepting group to another might be through "tRNA gene recruitment model" via point mutation in the anticodon [31]. Additionally, one important and interesting phenomenon is that CTG-decoding ser-tRNA was originated from a $tRNA^{Ser}$, which is consistent with the conclusions made from a recent phylogenetic study on $tRNA^{Ser}$s in *Candida* spp [11].

However, the exact relationships between some different tRNA species could not be defined directly from the mere analysis of the sequences. As shown in Fig. 2c, most tRNA species from *C. albicans* are evolved independently or their evolutionary positions are unsolved based on the sequence data provided. This is caused by the factor that tRNA genes are likely too small (about 75 nucleotides) to provide accurate information about their evolutionary history. It is possible that some mutations (forward and reverse) were accumulated in the pairs of tRNAs at variable sites to reach an equilibrium [32]. Clearly, these tRNA genes have a constant tertiary structure with a highly variable sequence.

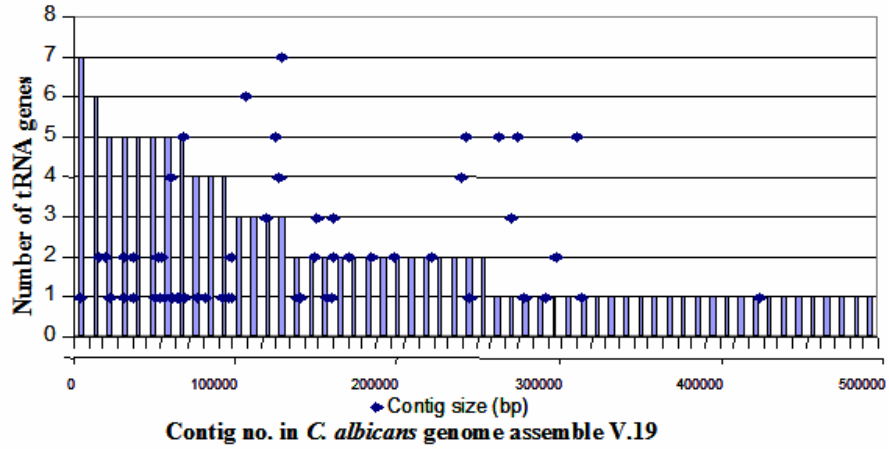
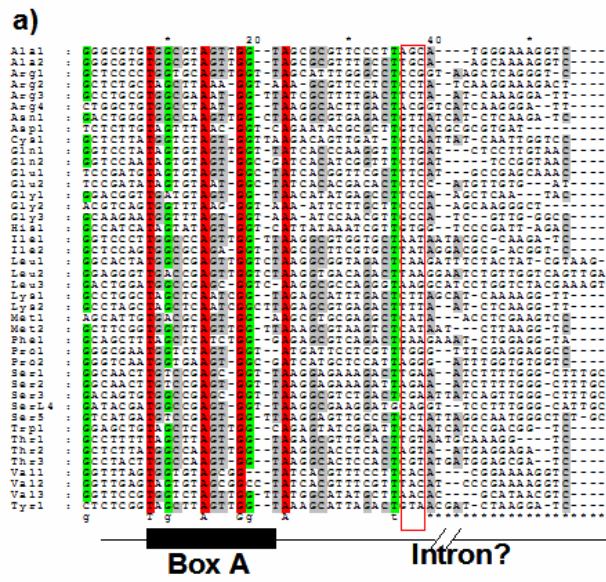
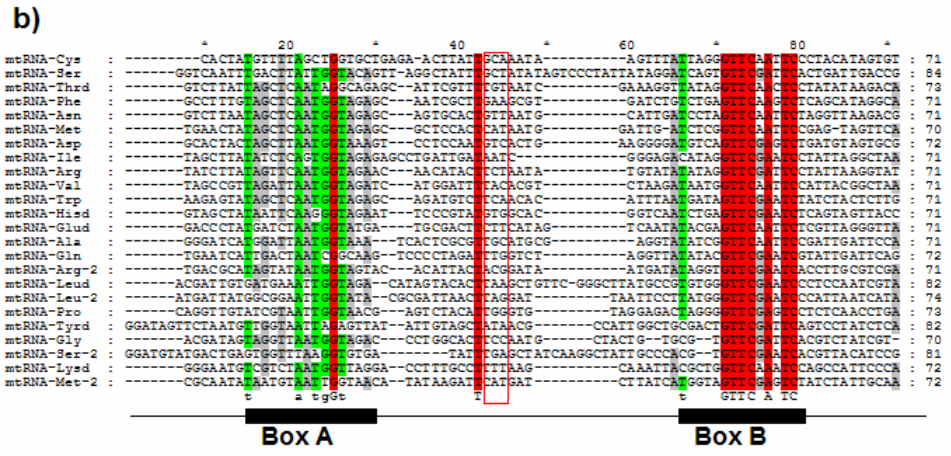


Fig. 1 Non-random distribution of the tRNA genes in the *C. albicans* genome. The vertical bars represent each contig containing the tRNA genes, and their corresponding contig sizes (bp) with the same number of the tRNA genes are shown as a dot with scale below the graph.





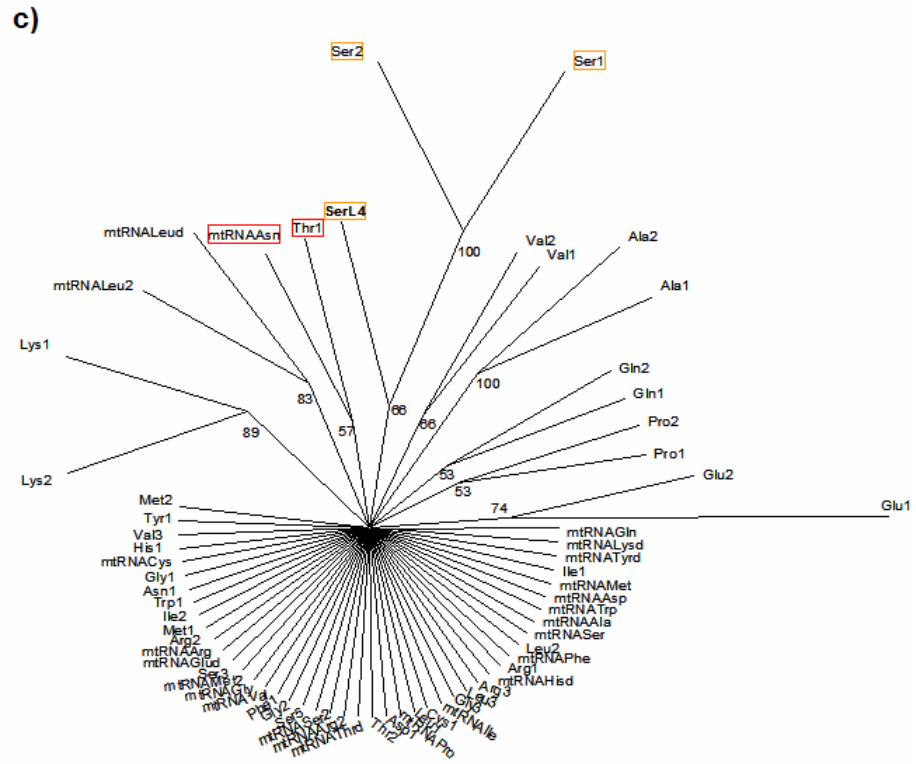


Fig. 2 tRNA sequences from *C. albicans* nucleus (a) and mitochondria (b), and their phylogeny (c). Alignments were constructed using Cluatal W and then manually adjusted. Dashes indicate gaps introduced to facilitate alignment. The blank box indicates the anticodon for each isoaccepting tRNA species. The consensus nucleotide acids are shown under each alignment as capital letters for all identical and small letters for more than 80% identity. The potential box A and box B, which are proposed to bind polymerase III, are schematically shown as solid black bars under each alignment. Intron positions for the potential insertion are also indicated. c). Phylogenetic relationships of tRNAs from *C. albicans*. The Bootstrap majority-rule consensus tree was constructed by using the maximum likelihood method for the tRNA sequences aligned using Clustal W and then manually adjusted. Bootstrap values from 100 replicates are shown.

3.4 *Correlation between tRNAs redundancy and codon preference*. Abundance of tRNA genes available in a cell has been well correlated with the frequency of the used codons [see [33,34]]. In comparison of the 132 tRNA gene species with codon usage estimated from the nonredundant 9259 ORFs (Fig. 4a and b), it is showed that the copy numbers of tRNA gene are positively correlated with overall codon frequency (r =0.758) (Fig. 4c). The correlation pattern between gene copy numbers and codon usage in *C. albicans* is in general remarkably similar to that exhibited for *S. cerevisiae* [9], and also is very consistent with the previous studies on many other organisms [35]. This suggests that increases of tRNA gene copy number is one of the main mechanisms to cope with the requirements of decoding capacity. Unexpectedly, however, the most preferred codons, ATT and AAC, encoding asparagine, are not correlated with the copy number of GTT-decoding $tRNA^{Asn}$, indicating that the encoding of the highly asparagine-rich proteins in *C. albicans* is achieved through increase of individual transcription rates and/or post-transcriptional regulatory mechanisms, which are different from the one employed by other tRNA genes [36]. As some of the asparagine-rich proteins are immune targets [37,38], it is worthwhile to study their biological functions further in this human pathogen. Once ignoring the point biased from general trend in the correlation analysis, it would display a higher correlation (r = 0.83) between the codon preference and the copy numbers of each tRNA gene (Fig. 4c). In consistence with previous reports on the positive correlation between intracellular concentration of tRNA isoacceptor and the

frequency of the synonymous codons [33,36,39], increase of tRNA genes copy number is likely a primary strategy for *C. albicans* in regulating intracellular tRNA concentration.

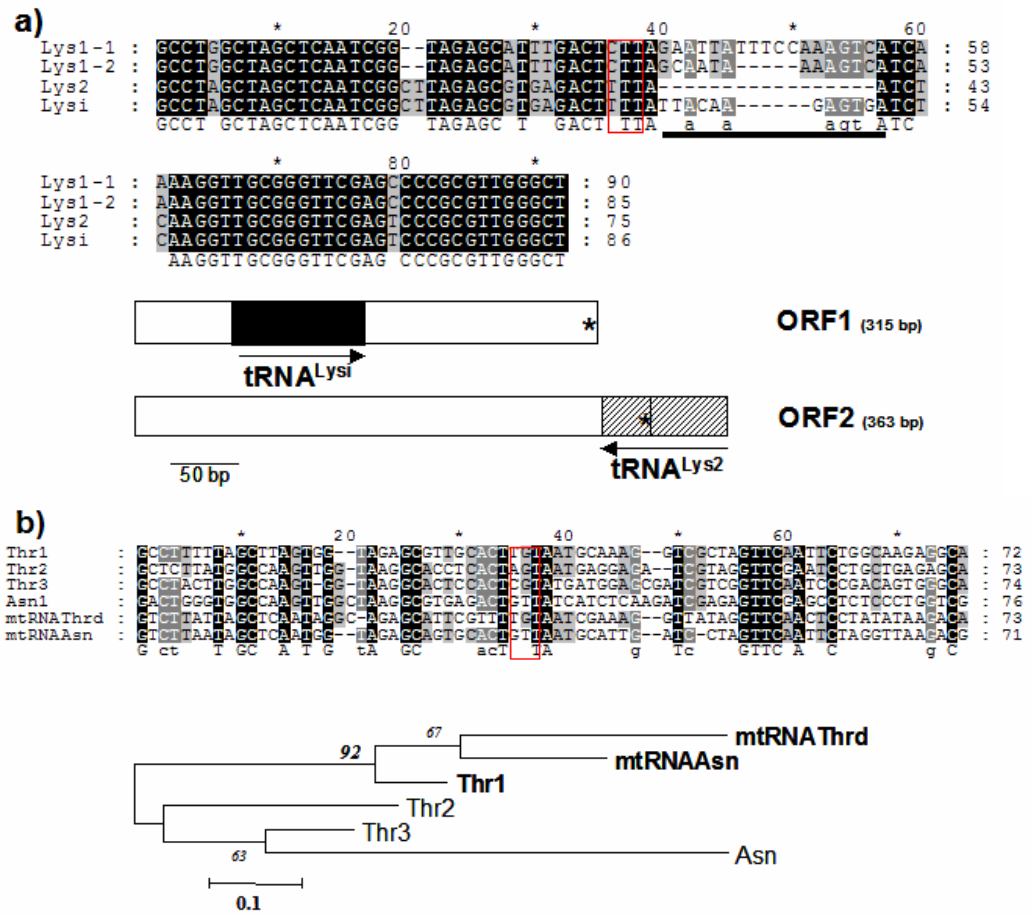


Fig. 3 Variation of tRNAs within the gene families and between the nuclear- and mitochondrial- encoding tRNAs. a). Sequence alignment of $tRNA^{Lys}$ with introns using the program Clustal W. The black and shadowed boxes represent conserved nucleotide acids. The black bar represents intron positions. The schematic shows that the $tRNA^{Lysi}$ gene is incorporated into ORF1 (315 bp) and partial $tRNA^{Lys2}$ gene is overlapped with ORF2 (363 bp). b). Sequence alignment and phylogenetic analysis of the nuclear- and mitochondrial encoding $tRNA^{Thr}$s and $tRNA^{Asn}$s. The black and shadowed boxes are conservative nucleotide acids. The phylogenetic tree was drawn by the Neighbour-joining method from a Clustal W sequence alignment shown as in the above. Bootstrap values supporting the branches from 100 replicates are shown at the note of each branch. The scale bar represents the number of substitutions per 100 sites for a unit branch length.

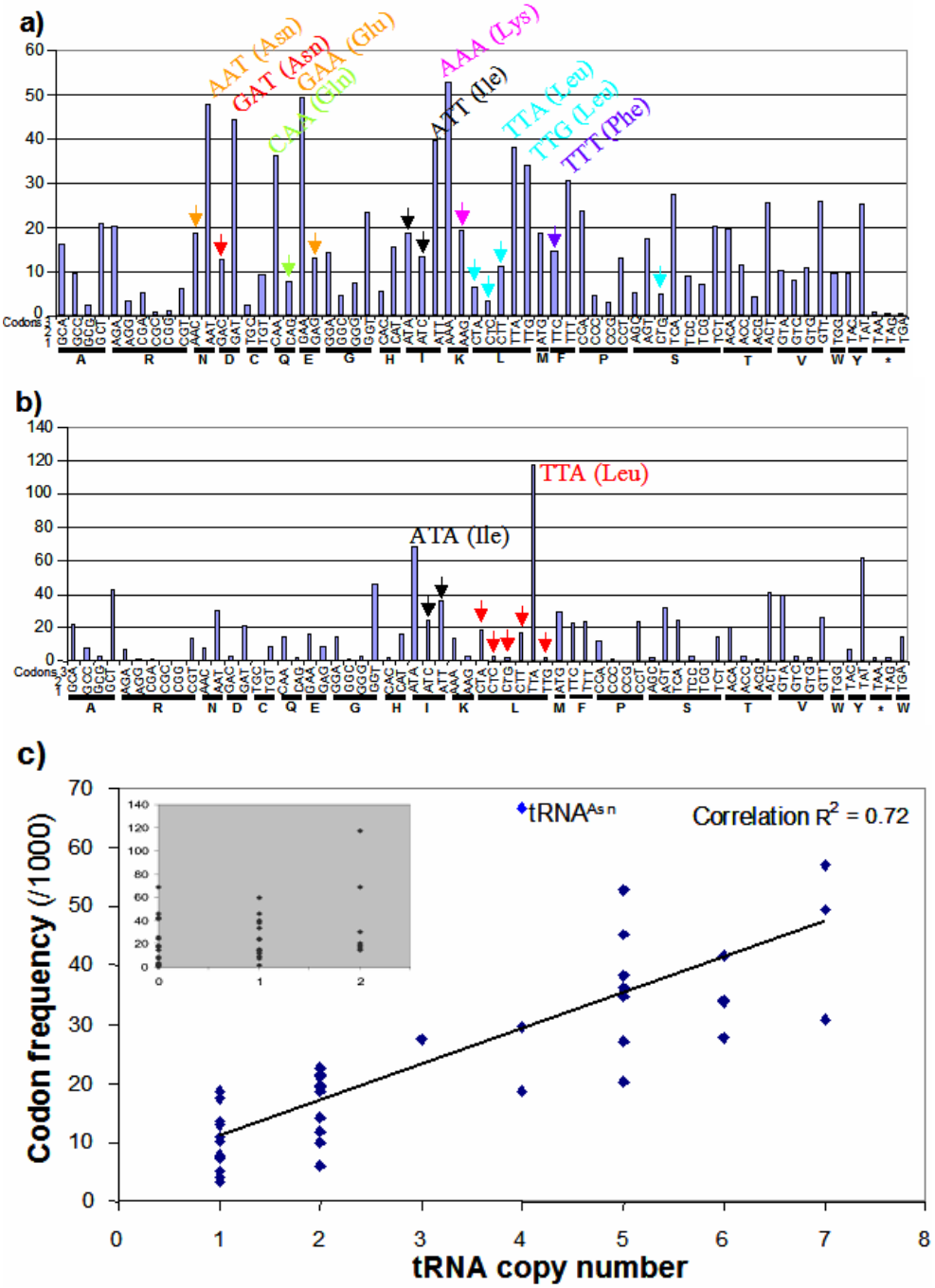


Fig. 4 Codon usage in *C. albicans* nucleus (a) and mitochondria (b), and its relationship with tRNA gene (c). Number of codons per thousand bases is shown as vertical bars. The top high-frequency of codons is indicated by arrows with different colors. The amino acids coded by synonymous codon triplets are shown in single letter code. c). Relationship between the copy number of isoaccepting tRNA species (x-axis) and codon frequency (y-axis) of *C. albicans*. When tRNAs were predicted to decode more than one codon, codon frequencies were calculated for rankings. Correlation r = 0.72. One exception, codons AAT and AAC, is highlighted to be marked as $tRNA^{Asn}$. The

graph in upper left corner represents the relationship between the copy number of mitochondrial-encoding tRNAs and codon usage of the mitochondrial genome.

The 24 unique tRNA genes from mitochondria (AF285361) decode most codons for all amino acids, whereas some codons are not decoded by any mitochondrial tRNA genes (Fig. 4c). Mitochondrial specific wobble rule may reduce the number of required anticodons substantially, thus providing a connection between genetic code and base paring stability of modified and unmodified RNA bases [40-42]. However, it could not be ruled out that mitochondria employs nuclear-encoding tRNAs to expand its decoding ability through importing the nuclear tRNAs into mitochondria as a compensation for the absent species as shown from the previous reports [43,44]. Therefore, the mitochondrial genome might use the nuclear-encoding tRNA genes to replace the completely lost tRNA genes (Fig. 4c). As shown in Fig. 4, unusual features of this genome include a complete missing of codons CGC, CCG, TCG, and TGG, as expected for A+T-rich genome, and all lost codons are G+C rich. Also, there is a substitution of stop codon TGA for TGG as the anticodon for $tRNA^{Trp}$. Unusual Leu-rich proteins for the encoding genes and the non-intron maintenance in the tRNA genes were also observed. Since the reassignment of codons in mitochondria occurs frequently [42] and its genetic codon capacity may be enhanced using specific wobble rules [41], the import of nuclear-encoding tRNA may not be the only solution to enhance its decoding ability.

In summary, the nearly completed sequences of the *C. albicans* genome greatly expedite molecular biology study on this human pathogenic yeast, thus contributing to the overall understanding of genome structure and comparative biology/evolution of this pathogen and its pathogenesis. Consequently, this has led to the identification of 132 tRNA genes with their key cellular functions in cellular protein synthesis. Specific gene duplications, interorganellar gene transfer, and asparagines-rich proteome were also observed from the *C. albicans* tRNA genes. Additionally, there was a positive correlation between *C. albicans* codon usage and tRNAs gene copy number. Although the current report is based on the genome data of *C. albicans* available now, the number of tRNA species in nucleus and mitochondria is amazingly the same as observed in the budding yeast *S. cerevisiae* [8,9]. Hence, this report may cover the complete set of tRNA genes from this human pathogen in terms of tRNA species and decoding capacity, even though the total number of tRNA genes from the *C. albicans* haploid genome is less than that in *S. cerevisiae*. As research on tRNAs has been increasingly engrossed [45], further studies will provide us more details about tRNA evolution and codon reassignment in this human pathogenic yeast.

Acknowledgments This work was in part supported by the Functional Genomics Program from the NRC GHI and NSERC Visiting Fellowships in Canadian Government Laboratories from Natural Sciences and Engineering Research Council of Canada to JF and FS. We acknowledge help and advice received from Drs. M.S. Whiteway, C. Lee, Z. Yu and C. Bachewich (BRI, N.R.C., Canada). Special thanks are due to M. Slater for very useful comments on the manuscript. DNA sequencing was performed at the Biotechnology Research Institute of N.R.C., Canada.

References

[1] L. P. Lim, M. E. Glasner, S. Yekta, C. B. Burge, and D. P. Bartel, Science 299, 1540 (2003).
[2] E. G. Moss and R. S. Poethig, Curr. Biol. 12, R688-R690 (2002).
[3] J. S. Mattick, EMBO Rep. 2, 986 (2001).
[4] A. Fire, S. Xu, M. K. Montgomery, S. A. Kostas, S. E. Driver, and C. C. Mello, Nature 391, 806 (1998).
[5] C. L. Liquori, K. Ricker, M. L. Moseley, J. F. Jacobsen, W. Kress, S. L. Naylor, J. W. Day, and L. P. Ranum, Science 293, 864 (2001).
[6] F. C. Odds, Candida and Candidosis, (London: Bailliere Tindall, 1988).
[7] T. M. Lowe and S. R. Eddy, Nucleic Acids Res. 25, 955 (1997).
[8] J. Hani and H. Feldmann, Nucleic Acids Res. 26, 689 (1998).
[9] R. Percudani, A. Pavesi, and S. Ottonello, J. Mol. Biol. 268, 322 (1997).
[10] C. Guthrie and J. Abelson, "Organization and expression of tRNA genes in Saccharomyces cerevisiae.," The Molecular Biology of the Yeast Saccharomyces: Metabolism and Gene Expression., J. Strathern, E. Jones, and J. Broach, eds., (Cold Spring Harbor Laboratory, Cold Spring Harbor, New York, 1982).

[11] S. E. Massey, G. Moura, P. Beltrao, R. Almeida, J. R. Garey, M. F. Tuite, and M. A. Santos, Genome Res. 13, 544 (2003).
[12] Y. Kawaguchi, H. Honda, J. Taniguchi-Morimura, and S. Iwasaki, Nature 341, 164 (1989).
[13] T. Suzuki, T. Ueda, T. Ohama, S. Osawa, and K. Watanabe, Nucleic Acids Res. 21, 356 (1993).
[14] J. D. Thompson, D. G. Higgins, and T. J. Gibson, Nucleic Acids Res. 22, 4673 (1994).
[15] J. Felsenstein, PHYLIP (Phylogeny Inference Package) Version 3.572, (Distributed over the World Wide Web, Seattle, USA, 1995).
[16] S. Kumar, K. Tamura, I. B. Jakobsen, and M. Nei, Bioinformatics. 17, 1244 (2001).
[17] R. D. Page, Comput. Appl Biosci. 12, 357 (1996).
[18] C. d'Enfert, S. Goyard, S. Rodriguez-Arnaveilhe, L. Frangeul, L. Jones, F. Tekaia, O. Bader, A. Albrecht, L. Castillo, A. Dominguez, J. F. Ernst, C. Fradin, C. Gaillardin, S. Garcia-Sanchez, P. de Groot, B. Hube, F. M. Klis, S. Krishnamurthy, D. Kunze, M. C. Lopez, A. Mavor, N. Martin, I. Moszer, D. Onesime, M. J. Perez, R. Sentandreu, E. Valentin, and A. J. Brown, Nucleic Acids Res. 33, D353-D357 (2005).
[19] F. H. Crick, J. Mol. Biol. 19, 548 (1966).
[20] D. C. Jaikaran, C. H. Collins, and A. M. MacMillan, J. Biol. Chem. 277, 37624 (2002).
[21] P. Medstrand, L. N. van de Lagemaat, and D. L. Mager, Genome Res. 12, 1483 (2002).
[22] P. F. Johnson and J. Abelson, Nature 302, 681 (1983).
[23] M. C. Strobel and J. Abelson, Mol. Cell Biol. 6, 2663 (1986).
[24] V. M. Perreau, M. A. Santos, and M. F. Tuite, Mol. Microbiol. 25, 229 (1997).
[25] M. A. Santos, G. Keith, and M. F. Tuite, EMBO J. 12, 607 (1993).
[26] V. I. Lim, J. Mol. Biol. 252, 277 (1995).
[27] A. Sidow, Curr. Opin. Genet. Dev. 6, 715 (1996).
[28] M. P. Cummings, J. M. Nugent, R. G. Olmstead, and J. D. Palmer, Curr. Genet. 43, 131 (2003).
[29] J. D. Palmer, K. L. Adams, Y. Cho, C. L. Parkinson, Y. L. Qiu, and K. Song, Proc. Natl. Acad. Sci. U. S. A 97, 6960 (2000).
[30] Y. Kobayashi, V. Knoop, H. Fukuzawa, A. Brennicke, and K. Ohyama, Mol. Gen. Genet. 256, 589 (1997).
[31] M. E. Saks, J. R. Sampson, and J. Abelson, Science 279, 1665 (1998).
[32] R. Holmquist, T. H. Jukes, and S. Pangburn, J. Mol. Biol. 78, 91 (1973).
[33] T. Ikemura, "Correlation between codon usage and tRNA content in microorganisms.," Transfer RNA in protein synthesis., D. L. Hatfield, B. J. Lee, and Pirtle.R.M., eds., (CRC Press, Boca Raton, 1992), 87-111.
[34] P. M. Sharp and W. H. Li, J. Mol. Evol. 24, 28 (1986).
[35] C. Marck and H. Grosjean, RNA. 8, 1189 (2002).
[36] H. Dong, L. Nilsson, and C. G. Kurland, J. Mol. Biol. 260, 649 (1996).
[37] K. Zhang, H. Fujioka, C. A. Lobo, D. Kitayaporn, M. Aikawa, and N. Kumar, Parasitol. Res. 85, 956 (1999).
[38] D. A. Fidock, E. Bottius, K. Brahimi, I. I. Moelans, M. Aikawa, R. N. Konings, U. Certa, P. Olafsson, T. Kaidoh, A. Asavanich, and et al., Mol. Biochem Parasitol. 64 , 219 (1994).
[39] T. Ikemura, J. Mol. Biol. 146, 1 (1981).
[40] T. H. Jukes and S. Osawa, Experientia 46, 1117 (1990).
[41] R. D. Knight, L. F. Landweber, and M. Yarus, J. Mol. Evol. 53, 299 (2001).
[42] S. Yokobori, T. Suzuki, and K. Watanabe, J. Mol. Evol. 53, 314 (2001).
[43] N. Entelis, O. Kolesnikova, H. Kazakova, I. Brandina, P. Kamenski, R. P. Martin, and I. Tarassov, Genet. Eng (N. Y.) 24, 191 (2002).
[44] N. Entelis, O. Kolesnikova, H. Kazakova, I. Brandina, P. Kamenski, R. P. Martin, and I. Tarassov, Genet. Eng (N. Y.) 24, 191 (2002).
[45] A. K. Hopper and E. M. Phizicky, Genes Dev. 17, 162 (2003).