Bramall, Ellis

Selective advantage of the tRNA I34 Deamination in Eukaryotes

Original Citation

Bramall, Ellis (2020) Selective advantage of the tRNA I34 Deamination in Eukaryotes. Masters thesis, University of Huddersfield.

University of Huddersfield

Selective advantage of the tRNA I34 Deamination in Eukaryotes

Author:
Ellis Bramall

Supervisor:
Dr. Martin Carr

A dissertation submitted in partial fulfilment of the requirements for the degree of Masters by Research

Department of Biological and Geographical Sciences

School of Applied Sciences

University of
HUDDERSFIELD
Inspiring tomorrow’s professionals

Copyright Statement

Declaration of Authorship

I, Ellis Bramall, declare that this dissertation titled, “Selective advantage of the tRNA I34 Deamination in Eukaryotes” and the work presented in it are my own. I confirm that:

- This work was done wholly or mainly while in candidature for a research degree at this University.

- Where any part of this dissertation has previously been submitted for a degree or any other qualification at this University or any other institution, this has been clearly stated.

- Where I have consulted the published work of others, this is always clearly attributed.

- I have acknowledged all main sources of help.

Signed:

Date: 18/09/2020

Acknowledgements

First and foremost a big thank you to Dr. Martin Carr for giving me this opportunity to carry out this study and contribute to your research. Working with you over these last two years has allowed me to discover the world of bioinformatics and my appreciation for it.

Secondly, I would like to thank Covid-19 for showing me that even in tough circumstances I have the ability to power through and get the job done. A special thanks to my mum who has constantly given me her love, support and guidance. I would also like to thank Layla, Georgia, Kristine, Anais, Charlie, Jack, Molly and Grace for collectively keeping me sane.

I also have to thank my grandparents, Marian and Ian, who not only provided the means and resources to undertake this research but have also been amazing and loving role models throughout my life.

Finally, I would like to thank my boyfriend Kyle for keeping me comfortable, motivated and fed through quarantine. Without you I would have never even considered undertaking this task, it was you who allowed me to see my own intelligence and worth.

Abstract

Introduction: A tRNA modification which deaminates adenosine into inosine at position 34 of tRNAs for the threonine, alanine, proline, serine, isoleucine, leucine, valine and arginine (TAPSILVR) amino acids within eukaryotes is carried out by an enzyme known as hetADAT. This deamination allows for certain codons, known as ADAT preferred codons, to be recognised by tRNA types which are usually non-complementary.

Aims: By looking at the patterns of codon usage within 33 different protists from a range of eukaryotic supergroups, an estimation into the origins of the TAPSILVR tRNA deamination in eukaryotes can be proposed. Secondly, by determining the codon usage of the protists, whether or not the TAPSILVR tRNA deamination allows for the selection of the ADAT preferred codons due to translational efficiency and/or accuracy can also be examined.

Methodology: Optimal codons were identified for each of the 33 protists with the use of CodonW and expression data provided by the NCBI SRA. Under a selection model, the optimal codons are those which correspond to the most abundant tRNA molecules, therefore, the number of genes present with each protist's genome was found for each tRNA. Certain patterns in the tRNA copy numbers can also show consistency with the presence of the I34 tRNA deamination. A34 tRNA molecules were modified to show the I34 deamination and screened for within available small tRNA data in order to provide evidence for the presence of the deamination.

Results: The codon usage of the majority of the protists along with their tRNA gene copy numbers showed consistency with not only the use of the I34 tRNA deamination but also for the selection of the ADAT preferred codons due to the deamination. tRNA molecules modified with the I34 deamination were found in species belonging to the eukaryotic supergroups of Amoebozoa, Opisthokonta, Excavata, Alveolata and Stramenopiles.

Conclusions: The presence of the I34 tRNA deamination has now been found in eukaryotic supergroups across the eukaryotic tree of life, whilst a selective pressure towards the use of ADAT preferred codons has been found to affect the codon usage bias of organisms from all of the eukaryotic supergroups researched. Therefore, this study has provided evidence that the origin of the I34 tRNA deamination carried out by hetADAT lies in the last common ancestor of the eukaryotes, where it most likely provided an advantage for the use of ADAT preferred codons.

List of Abbreviations and terms used

A34	Adenosine at the 34^{th} position
A37	Adenosine at the 37^{th} position
ADAT	Adenosine deaminase acting on tRNA
AT	Adenine and thymine nucleotide bases
BLAST	Basic Local Alignment Search Tool
CDS	Coding region of a gene
CUB	Coding usage bias
eToL	eukaryotic Tree of Life
ExPASy	Expert Protein Analysis System
FASTA	Text-based format for representing nucleotide/ amino acid sequences
Fop	Frequency of optimal codons
GC	Guanine and cytosine nucleotide bases
GC3s	Guanine or cytosine content at the third position of codons
HEG	Highly expressed genes
I34	Inosine at the 34^{th} position
I37	Inosine at the 37^{th} position
JGI	Joint Genome Institute
mRNA	Messenger Ribonucleic Acid
miRNA	Micro Ribonucleic Acid
Nc	Effective number of codons used
NCBI	National Center for Biotechnology Information
ncRNA	Non-Coding Ribonucleic Acid
ORF	Open Reading Frame
P-value	The probability that a difference occurred by random chance
SAM	Sequence Alignment Map
SAR	The supergroups of Stramenopiles, Alveolata and Rhizaria
SEQTK	Toolkit for processing sequences in the FASTA format
SILVR	The amino acids serine, isoleucine, leucine, valine and arginine
SMALT	DNA aligning tool
SRA	Sequence Read Archive
TadA	tRNA adenosine deaminase A
TAPSILVR	The amino acids threonine, alanine, proline serine, isoleucine, leucine, valine and arginine
tRNA	Transfer Ribonucleic Acid
WEG	Weakly expressed genes

Contents

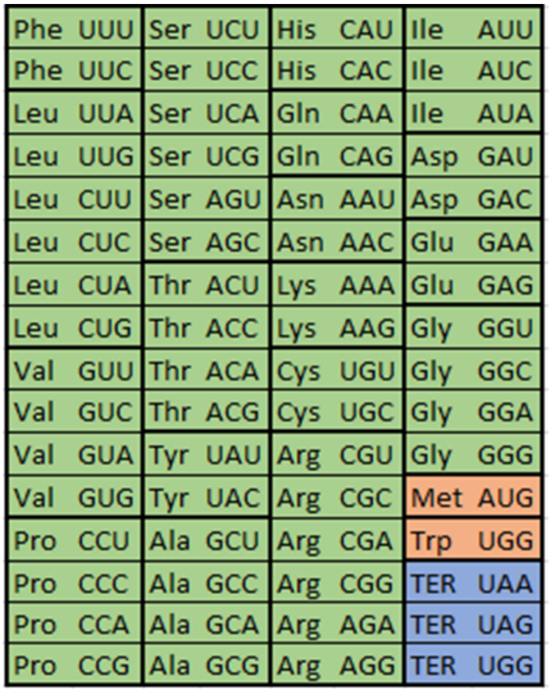
1.0 Introduction

1.1 Codon Usage Bias

In all known organisms DNA is used to transcribe RNA, which is translated into proteins made up of amino acid monomers. As there are 20 different amino acids, their usage is determined by codons, three nucleotide codes found in mRNA. There are 64 codons in total, however in the standard genetic code 61 of these codons code for amino acids, whilst the remaining three act as STOP codons. In other genetic codes this is not always the case, notably within certain ciliates the codons UAA and UAG, known as STOP codons within the standard genetic code, code for glycine when found at internal positions within genes (Jukes & Osawa, 1993).

1.1.1 Synonymous Codons

As there are 61 codons coding for 20 amino acids, the majority of amino acids are coded for by multiple codons, ranging from two to six codons each (Figure 1). Methionine and tryptophan are the only two amino acids which are coded for by a singular codon. Codons which code for the same amino acids are known to each other as synonymous.



Phe UUU	Ser UCU	His CAU	Ile AUU
Phe UUC	Ser UCC	His CAC	Ile AUC
Leu UUA	Ser UCA	Gln CAA	Ile AUA
Leu UUG	Ser UCG	Gln CAG	Asp GAU
Leu CUU	Ser AGU	Asn AAU	Asp GAC
Leu CUC	Ser AGC	Asn AAC	Glu GAA
Leu CUA	Thr ACU	Lys AAA	Glu GAG
Leu CUG	Thr ACC	Lys AAG	Gly GGU
Val GUU	Thr ACA	Cys UGU	Gly GGC
Val GUC	Thr ACG	Cys UGC	Gly GGA
Val GUA	Tyr UAU	Arg CGU	Gly GGG
Val GUG	Tyr UAC	Arg CGC	Met AUG
Pro CCU	Ala GCU	Arg CGA	Trp UGG
Pro CCC	Ala GCC	Arg CGG	TER UAA
Pro CCA	Ala GCA	Arg AGA	TER UAG
Pro CCG	Ala GCG	Arg AGG	TER UGG

Figure 1 | The Standard Genetic Code This figure shows the amino acids and the codons which correspond to them. Codons in blue represent synonymous STOP codons. Codons in orange are non-synonymous. Codons in green are synonymous.

The first two bases of synonymous codons, excluding those for serine, leucine and arginine, are conserved, whilst the third base differs. Therefore, a case could be made that for the purpose of translation, the third base, in certain synonymous codons, is redundant. According to The Wobble Hypothesis, this redundancy allows for non-Watson-Crick base pairings to occur between the third base of a codon and the first base of the tRNA anticodon (Crick, 1966). It is this ‘wobble’ which allows for the pairing of uracil and guanine. The ‘wobble’ also allows for the modified adenine base, inosine, to pair with adenine, cytosine and uracil (Gerber, 1999).

Although synonymous codons code for the same amino acids, they are not used in equal frequencies; depending on the organism, certain synonymous codons are utilised more than others (Grantham *et al*., 1980). This unequal distribution of synonymous codons is known as codon usage bias (CUB) (Comeron and Aguadé, 1998). CUB not only varies between species but it also differs between genes present within the same genome (Grantham *et al*., 1980). Within the AT rich *Dictyostelium discoideum* genome for example there is a large degree of codon usage variation present within silent sites of genes, with the GC content ranging from under 10% to above 30% (Sharp & Devine, 1989). The length of genes has been shown to positively correlate with codon usage in both *Escherichia coli* and *Gnetum luofuense* (Moriyama, 1998; Deng *et al*., 2020); whereas a negative correlation was observed in *Saccharomyces cerevisiae* and *Drosophila melanogaster* (Moriyama, 1998). Although studies have suggested that the main factor which determines codon usage within genes is their expression level (Zhou *et al*., 2016). This correlation has been observed in many organisms including *Cuscuta australis*, *Drosophila melanogaster* and *Pseudomonas aeruginosa.* (Liu *et al*., 2020; Duret & Mouchiroud,1999; Gupta & Ghosh, 2001).

To determine the overall CUB of a gene the ‘effective number of codons’ (Nc) of the gene is measured (Wright, 1990). This value can range from 20, where only one codon is used per amino acid, to 61 were all codons are present at equal frequency within the gene. There are many factors which could possibly influence CUB, the three main mechanisms being mutational pressure, genetic drift and natural selection (Figuet *et al*., 2014; Machado *et al*., 2019). The CUB for each gene is due to a combination between these mechanisms and therefore, rarely does a single mechanism influence CUB independently.

1.1.2 Mutation Pressure

When DNA is replicated there is a chance that mutations can occur, differences in local mutation rates therefore, can result in base composition changes across genomes. However, due to other influences on CUB, mutational pressure is hard to observe except when its biases are extreme. In these extreme cases, mutational pressure has forced whole genomes to shift to either extreme GC or extreme AT content (Eichinger *et al*., 2005, Scala, 2002). In prokaryotic genomes GC content has been shown to range from 25% to 75% (Musto *et al.*, 2006), highlighting a strong mutational bias influencing codon usage. This bias can also be seen in protists, for instance within the genome of *Dictyostelium discoideum* which has a GC content of 25% (Eichinger, 2005).

Due to the redundancy of codons, the best way to measure mutational drift within an organism is to find the GC content at the third position of codons (GC3s). However, if a genes' codon usage is subjected to a selective force stronger than that of mutational pressure, the GC3s value of the gene would no longer be neutral, and therefore, a weak indicator of mutational pressure. This could be somewhat resolved by looking at GC content in non-coding DNA as these regions are more likely to remain neutral from natural selection.

1.1.3 Selection

Natural selection can act upon organisms' genes with the use of codons that provide a selective advantage when utilised instead of their synonymous counterparts; these codons are referred to as "optimal codons". For many species, the use of optimal codons within genes is strongly correlated with expression levels and the abundance of corresponding tRNA molecules (Goy and Gautier, 1982; Sharp *et al*., 2010). When first conceived, optimal codons were considered to be those that complemented the most abundant tRNA isotype (Ikemura, 1981). However, possibly due to the difficulty of quantifying tRNA levels, an alternative method to identify optimal codons was implemented by identifying those codons which appeared more frequently in highly expressed genes (Akashi and Eyre-Walker, 1998).

As expression data is not always available, a program called CodonW (Pedan, 1999) identifies codons found in higher frequencies in biased genes compared to genes with little bias; based upon axis 1 of a correspondence analysis. However, this method works under the assumption that mutational pressure is a minor factor on CUB (Pedan, 1999). The frequency of optimal codons

(F_{op}) of a gene can be calculated by dividing the number of optimal codons found in the gene by the total amount of codons in said gene (Ikemura, 1981).

There are two mechanisms through which selection may operate, though they are not mutually exclusive: translational efficiency and translational accuracy.

1.1.4 Translational Efficiency and Translational Accuracy

As there are a limited number of ribosomes in a cell, when translating genes at high rates, their low quantity can become a limiting factor (Metzl-Raz *et al*., 2020). This can be minimised by the utilisation of optimal codons which speed up the translation process faster than their synonymous counterparts (Sørensen & Pedersen, 1991). Therefore, if selection was operating under translational efficiency, the degree of selective pressure would be higher on the codon usage in highly expressed genes than that of weakly expressed genes. In unicellular organisms, like the protists examined in this study, growth rate and the rate of protein synthesis are positively correlated (Pakula, 2005). A high growth rate indicates a high chance of reproductive success, emphasising the advantage that selection due to translational efficiency provides for protists.

Selection also may operate in order to maximise the accuracy of translation. If a protein is subject to amino acid misincorporation during translation there is a chance that its function could be lost. This indicates a need for selection to increase translational accuracy. Precup and Parker (1987) showed that the use of optimal codons within *E. coli* genes reduced amino acid misincorporation 10-fold, compared to that of non-optimal synonymous codons. Translational accuracy is of greater importance to the protein coding domains of a gene than their non-domain encoding regions of genes. Therefore, when under a selection via translational accuracy model optimal codons are found at higher frequencies in the protein coding domains when compared to the non-domains, this can be seen within *Drosophila* transcription factor encoding genes (Akashi, 1994).

1.1.5 tRNA Abundance

As previously mentioned, optimal codons are usually those recognised by the most abundant tRNA. It is expected that tRNA abundance is determined by the amount of tRNA genes present for each type in the genome, known as their copy number (Iben & Maraia, 2012). Thus, in species where selection influences codon usage, the copy number for each tRNA within a genome determines the optimal codons; therefore, the tRNA copy number is likely to be an important factor in CUB.

The copy number for each tRNA varies between species and differs greatly between kingdoms (Marck & Grosjean, 2002). Prokaryotes normally have around 20-25 different tRNAs, whilst eukaryotes can have up to 40-45 different tRNAs in their genomes (Grosjean *et al.*, 2009). For each case however, the number of different tRNA is always lower than the number of different codons used in the genome (Iben & Maraia, 2012). Therefore, in order for each codon to be recognised some tRNAs have to recognise multiple codons, this is achieved via the wobble position.

1.2 A-to-I modification

Adenosine deaminase acting on tRNAs (ADAT) are a family of enzymes that deaminate adenosine to inosine in tRNA. As previously mentioned inosine on the first base of tRNA anticodons (I34) can pair, via ‘the wobble’, with adenine, cytosine and uracil (Figure 2) (Crick, 1966). This suggests that an anticodon with inosine at the wobble position has the ability to correspond with three different codons, potentially allowing a single species of tRNA to fulfil the role of three different tRNA species. ADAT enzymes are present in all domains of life (Zinshteyn & Nishikura, 2009), however ADAT enzymes that convert adenosine present on the first base of tRNA anticodons (A34) to I34 are only present in Bacteria and Eukaryota (Gerber, 1999).

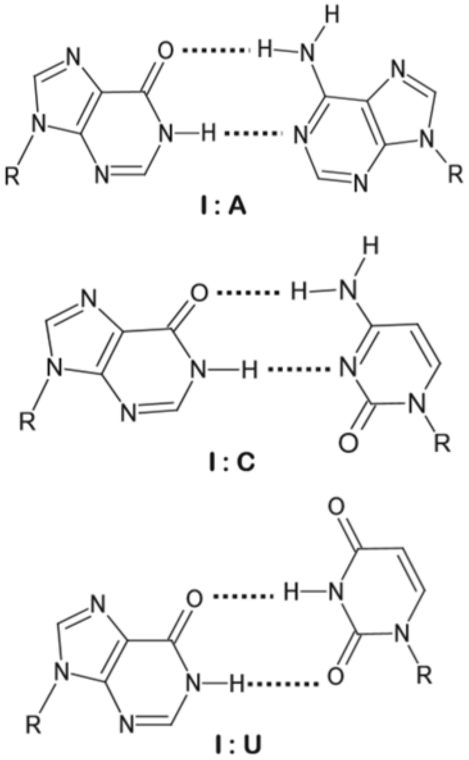


Figure 2 | Schematic representation of inosine wobble base pairings to adenine, cytosine and uracil. (Crick, 1966)

In Bacteria, the A34-to-I34 modification is catalysed by tRNA adenosine deaminase A (TadA), which carries out this interaction on only tRNA^{Arg} (Torres, 2014). In eukaryotes, however, the ADAT enzyme, a heterodimer comprised of ADAT2/ADAT3 (hetADAT), modifies base 34 in eight different tRNA types: tRNA^{Thr}, tRNA^{Ala}, tRNA^{Pro}, tRNA^{Ser}, tRNA^{Ile}, tRNA^{Leu}, tRNA^{Val} and tRNA^{Arg} (Boccaletto *et al*., 2017). The eight amino acids that the tRNAs correspond to are known collectively as TAPSILVR amino acids. Notably, tRNA^{Gly} is not listed despite the presence of the other 4-box tRNAs; this is due to the absence of A34 tRNA^{Gly} in genomes due to its unstable nature (Grosjean *et al*., 2009).

A third ADAT enzyme has also been found in eukaryotes, ADAT1. Instead of A34, ADAT1 modifies A37 into I37 in $tRNA^{Ala}$ only. It has recently been suggested in Torres (2014) that the modification of A37 by ADAT1 is independent of the presence of I34. The evidence does suggest that the I34 modification can and does occur independently of the I37 modification (Maas, Gerber & Rich, 1999), although the supporting literature provided in the review does not provide evidence that the I37 modification is independent of the I34 modification. Therefore, it is unclear if ADAT1's interaction depends on the interaction of hetADAT. ADAT1's interaction could also depend on the presence of I34, thereby only modifying the products of hetADAT activity.

The presence of A34 containing tRNAs is strongly correlated with the appearance of ADAT enzymes (Novoa *et al*., 2012). Hence, within eukaryotes, where hetADAT corresponds to many tRNA types, the amount of $tRNA_{ANN}$ types has been found to be much greater than Archaea or Bacteria (Chan & Lowe, 2009). In fact, in Archaea, where ADAT enzymes are absent, $tRNA_{ANN}$ types are not found (Marck & Grosjean, 2002). In bacteria, the only $tRNA_{ANN}$ found is $tRNA^{Arg}_{ACG}$, which is the only tRNA molecule deaminated by TadA (Wald & Margalit, 2014). This poses the question whether the appearance of ADAT enzymes in eukaryotes and bacteria drove selection for $tRNA_{ANN}$ types or whether the appearance of $tRNA_{ANN}$ types has led to the adaptation of the ADAT enzymes. Nonetheless, the change in the tRNA diversity between the three domains of life is directly correlated to the ADAT enzymes ability to potentially impact codon usage.

As stated, the deamination of adenosine to inosine at the wobble position of a tRNA suggests an increase in the number of codons the tRNA can recognise. However, due to the complicated rules of anticodon:codon pairing, there are some base pairings that are less efficient than others. Computer simulations have shown that A34 tRNAs are capable of reading C, A or G ending codons (Lim, 1995), although these pairings have been shown to decrease translational efficiency in the ribosome (Lim, 2001). The pairing of A34 tRNAs with C-ending codons within ribosomes also results in a destabilising action between the P-site and the A-site, whereas the utilisation of I34 tRNA has been shown to reduce this effect (Lim, 1995). In fact, I34 tRNAs bind preferentially to C with an increase in translational efficiency (Torres *et al*., 2014). Although, I34 tRNAs can also read U and A ending codons, translating such codons has been shown to be inefficient (Percudani, 2001). However, the lack of efficiency is less important in weakly expressed genes however, so I:A and I:U binding could be tolerated within these genes.

It is clear that, when utilised, the I34 modification allows for a greater accuracy when reading C-ending codons than that of A34 tRNA. Additionally, since A34 tRNA isotypes are rarely seen in the same genome as their G34 counterparts (Torres *et al*., 2014), the evidence suggests that I34 tRNAs are utilised in order to read C-ending codons efficiently in genomes lacking the G34 tRNA isotypes (Grosjean *et al*., 2009). There is a certain advantage to both the efficiency and accuracy of translation when I34 tRNA as opposed to A34 tRNA are utilised to read C-ending codons, therefore these codons are known as “ADAT-preferred codons” (Torres, 2014). ADAT-preferred codons have recently been shown in Rafels-Ybern (2017) to be abundant within the genomes of organisms within the eukaryotic kingdoms of Metazoa, Plantae and Fungi and Protists. Therefore, as the use of ADAT-preferred codons is widespread throughout eukaryotes, the origin of TAPSILVR tRNA deamination most likely lies within the protists, or within their last common ancestor.

1.3 Eukaryotic Tree of Life

Although the most well-known eukaryotes are animals, plants and fungi, the majority of eukaryotes are single-celled protists (Del Campo *et al*., 2014). Protists are often overlooked in eukaryotic based studies (Sibbald & Archibald, 2017); however, since the vast amount of eukaryotic diversity is found within protists, their evolutionary importance cannot be ignored. Due to their diversity, the eukaryotic Tree of Life (eToL) has been divided into 'supergroups'. Over the years, the number of these supergroups has fluctuated repeatedly between five and eight as a result of the use of phylogenetics and the continuous discoveries of new lineages of eukaryotes (Simpson & Roger, 2002; Keeling *et al*., 2005). Along with the quantity of supergroups, the placement of each supergroup on the tree has also varied as the exact position of the root has yet to be determined (Burki *et al*., 2019). Due to this for the purpose of this study, the interpretation of the eToL by Brown (2018) which utilises the Opimoda-Diphoda root, will be used. The Opimoda-Diphoda root segregates the super groups of Amoebozoa, Opisthokonta and Apusozoans from the supergroups of Excavata, SARs and Archaeplastida. This study will look at 33 organisms from all eukaryotic supergroups (Figure 3), excluding Archaeplastida, due to time constraints.

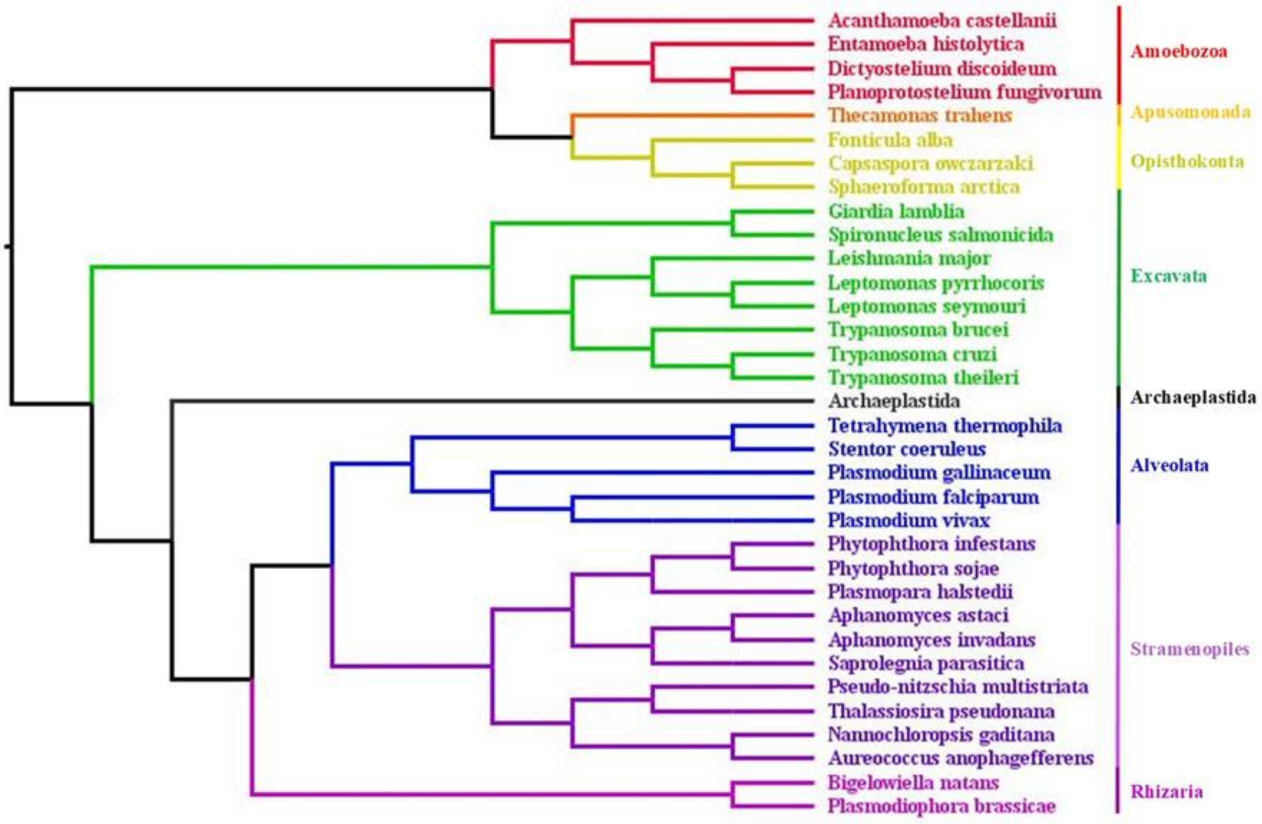


Figure 3 | Phylogeny of 33 protists showing the evolutionary relationships between the eukaryotic supergroups, based upon the phylogeny presented by Brown (2018) Constructed with the use of FigTree v. 1.4.3 (Rambaut, 2009).

1.3.1 Amoebozoans, Opisthokonta and Apusozoans

Amoebozoans are often characterised by the morphological trait of possessing pseudopodia or lamellipodia which can aid in locomotion or feeding (Lahr, Bergmann & Lopes, 2008). However, there is a noticeable discord between the molecular data and the morphological data of amoebozoans when determining phylogeny (Lara *et al*., 2008). Although both types of data have been utilised in unison to form groupings of amoebozoans (Bapteste *et al*., 2002), the relationships between these groups are not so easily resolved (Pawlowski & Burki, 2009). This study looks at four species of amoebozoans from different lineages in order to increase the possibility of closing this gap in knowledge. These species are: *Acanthamoeba castellanii,* an opportunistic pathogen from the class Discosea (Shabardina *et al*., 2018), *Entamoeba histolytica*, an Archamoebae parasite which lacks mitochondria (Loftus *et al*., 2005) and *Dictyostelium discoideum* and *Planoprotostelium fungivorum*, which are two slime molds that have the ability to aggregate into larger 'organisms' exhibiting a form of multicellularity (Eichinger *et al*., 2005; Bapteste *et al*., 2002).

The opisthokonts are probably the most well-known grouping of eukaryotes, consisting of both animals and fungi, however the group also includes many protists, defined by a posterior flagellum. Recent analysis of codon usage in three opisthokont protists pointed towards the utilisation of the A-to-I deamination in translation (Southworth *et al*., 2018). Previous studies, which found evidence for hetADATs within metazoan and plant genomes, speculated that this modification was important in the evolution of multicellularity (Rafels-Ybern *et al*., 2017; Porath *et al*., 2017). Therefore, Southworth (2018) questions this idea by potentially revealing the presence and utilization of hetADATs prior to multicellularity in eukaryotes. This study attempts to expand upon the evidence found by Southworth (2018) by examining *Capsaspora owczarzaki*, which will be examined in greater depth, and similar organisms: *Sphaeroforma arctica*, which possess a unique capacity to grow into multi-nucleated cells without displaying any of the usual growth stages (Suga & Ruiz-Trillo, 2013), and *Fonticula alba*, a slime mold which forms aggregative multicellular fruiting bodies (Brown, Spiegel & Silberman, 2009). Each of these organisms share certain insights behind the mechanisms that led to true multicellularity.

Apusozoans are grouped together by their use of flagella and the presence of an organic shell found under their dorsal surface (Cavalier-Smith & Chao, 2010). Although Apusozoa is often grouped with Opisthokonta and Breviatea in the clade Obazoa, it is not recognised as a eukaryotic supergroup. However, the Apusozoan *T. trahens* has been included within this study in order to not neglect this minor phylum.

A previous study that focused on transposable elements in protists found evidence consistent with the presence of the A-to-I deamination by hetADATs within *Fonticula alba* and *Thecamonas trahens* (Bramall, 2019). The inclusion of these organisms within this study will serve to confirm or dispute this evidence. If confirmed, this again opens up the possibility of hetADATs strongly influencing codon usage and therefore moulding entire genomes prior to multicellularity. As the extent of hetADATs influence over eukaryotic codon usage can only be fully realised if both sides of the eToL are looked at, this study also examines protists within supergroups SARs and Excavata.

1.3.2 Excavata, Stramenopiles, Alveolata and Rhizaria

Excavates are protists that possess feeding grooves on one of their sides and they are also known to cause numerous parasitic diseases in animals and humans (Ortiz & Solari, 2019). The Excavata examined in this study are: *Giardia lamblia*, a well-known intestinal parasite responsible for Giardiasis (Einarsson, Ma'ayeh & Svärd, 2016), *Spironucleus salmonicida*, a fish parasite (Xu *et al*., 2014) and *Leishmania major*, a pathogen associated with cutaneous leishmaniasis (Ivens, 2005). Alongside *T. brucei*, *T. cruzi* and *T. theileri*, members of the *Trypanosoma* genus which cause many parasitic infections, most commonly African Trypanosomiasis (Berriman, 2005; El-Sayed, 2005), and finally *L. pyrrhocoris* and *L. seymouri*, members of the Leptomonas genus (Flegontov *et al*., 2016; Kraeva *et al*., 2015).

The supergroup of SARs is made up of Stramenopiles, Alveolata and Rhizaria. Organisms within the supergroup Stramenopiles are categorised morphologically by the presence of a textured flagellum. Most Stramenopiles are algae, however, they also include oomycetes like *Phytophthora infestans*, the cause of the Irish potato famine (Haas *et al*., 2009). To determine the presence of the A-to-I modification within Stramenopiles a large variety of organisms from the supergroup were studied: *Aphanomyces astaci*, *Aphanomyces invadans, Aureococcus anophagefferens*, *Nannochloropsis gaditana*, *Phytophthora infestans, Phytophthora sojae*, *Plasmopara halstedii*, *Pseudo-nitzschia multistriata, Saprolegnia parasitica* and *Thalassiosira pseudonana*.

Alveolata are an extremely diverse set of organisms known collectively by a few morphological similarities (Leander *et al*, 2003). Due to their diversity alveolates are separated into three morphological subgroups: apicomplexans, ciliates and dinoflagellates. This study includes two ciliates, *Tetrahymena thermophila* and *Stentor coeruleus*, and three apicomplexans from the *Plasmodium* genus: *P. gallinaceum*, *P. falciparum* and *P. vivax*. Rhizaria is said to be the least understood of the eukaryotic supergroups (Burki & Keeling, 2014); this could be due to the fact that no rhizarian organisms are known to affect humans directly. Due to this bias complete rhizarian genomes are rare, which limited this study to two rhizarians: *Bigelowiella natans* and *Plasmodiophora brassicae*.

1.4 Aims and Objectives

The overall aim to the research carried out within this study is to provide an estimation into the origins of the TAPSILVR I34 tRNA deamination in eukaryotes. This will be achieved by comparing the usage of the ADAT preferred codons within the protists to the abundance of the different tRNA genes present within their genomes. Due to the recent availability of genomes and transcriptomes, research can be conducted on a wide range of protists from nearly all eukaryotic supergroups. Therefore, by looking at the distribution of tRNA deamination across eukaryotes the potential origin of the deamination within eukaryotes can be determined.

Research has shown that the eukaryotic ADAT1 evolved through gene duplication and divergence into the subunits of hetADAT (ADAT2 and ADAT3) (Gerber, 1999). Therefore, if the presence of the A-to-I deamination was found in organisms of supergroups from both sides of the eukaryotic tree, it would suggest that this duplication occurred in a eukaryotic ancestor to both clades. The relationship between the ADAT1 enzyme and the hetADAT enzyme, along with their respective modifications, is yet to be fully realised. Therefore, within this study both the I34 and the I37 modifications on $tRNA^{Ala}$ genes will be compared to find any dependency between the two.

If evidence of the I34 deamination was found across the early eukaryotes observed in this study, the mechanisms behind the use of the ADAT preferred codons will be investigated in terms of natural selection. If the ADAT preferred codons were found in higher frequencies within the highly expressed genes of an organism when compared to its weakly expressed genes this would indicate that the usage of the ADAT preferred codons within the organism was due to selection via translational efficiency. Similarly, if the ADAT preferred codons were found in higher frequencies within the protein coding domains of the highly expressed genes of an organisms when compared to the non-coding domains, it would appear that the ADAT preferred codons were selected for in order to maximise translational accuracy within the organism.

This study will therefore be investigating three lines of research:

1. The estimation of the origin of the TAPSILVR I34 tRNA deamination in eukaryotes.
2. The determination of the forces behind the potential bias for the ADAT preferred codons.
3. The relationship between the modifications of ADAT1 and hetADAT.

2.0 Materials and Methods

2.1 Data Collection

The genome files for all 33 organisms investigated within this study were provided by EnsemblProtists, an online database housing genetic data for protists (Kersey et al., 2018). The annotated transcriptomes for each organism were also provided for by EnsemblProtists. Genomic data was stored as text files in FASTA format. Any viewing or manipulating of these genomic files was carried out using Notepad++ (version 7.8.8).

Transcriptomic expression data for all organisms in the study were provided by the NCBI Sequence Read Archive (SRA). Where possible ncRNA and miRNA transcriptomic data was also retrieved from the NCBI SRA. The accession numbers for each organisms' set of transcriptomic data can be found on Table 1, the choice of small RNA data for each organism is also shown.

Name	RNA data	ncRNA	miRNA
Acanthamoeba castellanii	SRX339674	-	-
Entamoeba histolytica	DRX099774	-	SRX219151
Dictyostelium discoideum	SRX7692151	-	SRX150810
Planoprotostelium fungivorum	SRX2638173	-	-
Thecamonas trahens	SRX096916	-	-
Fonticula alba	SRX183037	-	-
Capsaspora owczarzaki	SRX155797	-	ERX2681962
Sphaeroforma arctica	ERX2270207	-	ERX2678837
Giardia lamblia	SRX625517	-	SRX625517
Spironucleus salmonicida	SRX332235	-	-
Leishmania major	ERX190352	SRX541170	-
Leptomonas pyrrhocoris	SRX1044319	-	-
Leptomonas seymouri	SRX1046748	-	-
Trypanosoma brucei	SRX271982	SRX1071658	-
Trypanosoma cruzi	SRX574890	-	-
Trypanosoma theileri	ERX1819477	-	-
Tetrahymena thermophila	SRX5736565	-	-
Stentor coeruleus	SRX2366376	-	-
Plasmodium gallinaceum	SRX729933	-	-
Plasmodium falciparum	ERX3875738	SRX4668733	-
Plasmodium vivax	ERX1787321	-	-
Phytophthora infestans	SRX4277324	SRX759739	-
Phytophthora sojae	SRX4277319	-	SRX2268634
Plasmopara halstedii	ERX541772	-	-
Aphanomyces astaci	SRX236909	-	-
Aphanomyces invadans	SRX238128	-	-
Saprolegnia parasitica	SRX155938	-	-
Pseudo-nitzschia multistriata	SRX1070747	-	-
Thalassiosira pseudonana	SRX3397332	SRX553515	-
Nannochloropsis gaditana	SRX151683	-	-
Aureococcus anophagefferens	SRX4623563	-	-
Bigelowiella natans	DRX053601	-	-
Plasmodiophora brassicae	SRX4994054	-	-

Table 1 | Organisms and SRA accession numbers A list of all organisms and their transcriptomic data's SRA accession numbers, which correspond to the experiments in which the data was originally found.

2.2 Determining Gene Expression Levels

SMALT was used in order to map sequencing reads onto their respective annotated gene sequence (Hannes Ponstingl, Genome Research Ltd). For this to be done, each individual organism's CDS file was used to create an individual hash index in the form of two separate files, one with the extension .smi and the other with the extension .sma. The hash index for each organism was then loaded into memory and mapped via paired-end reads to the file containing the relevant organism's whole transcriptome. The output to this step was written onto a file with the extension .sam in a SAM output format. The lines of text which code for both the creation of the hash table and the .sam file are found at https://www.sanger.ac.uk/tool/smalt-0/.

Tablet was then used to tabulate the number of reads per CDS (Milne et al., 2013). As genes of longer length will have more reads mapped to them by SMALT than a gene of a shorter length with the same expression level expression was normalised by dividing the number of reads by the CDS length.

The genes showing zero expression were then removed from the data set. The top and bottom 5% expressed genes were then extracted from the organisms genomic file and placed in two .txt files respectively. This separation was performed with the use of the tool SEQTK (Li, 2013). The top and bottom 5% expressed genes were extracted in order to compare codon usage between the two extremes of expression. Additionally, the top 200 expressed genes were also extracted from the organisms genomic file using SEQTK. These steps were carried out for each of the 33 organisms investigated within this study.

2.3 Separation of Domains and Non-Domains

The top 200 expressed genes from each organism were each individually separated into protein coding domains and non-domains with the use of the NCBI protein annotation. Codons from the two distinct gene regions were placed into separate FASTA files. This was completed for all organisms investigated in this study, excluding the top 200 expressed genes of *B. natans*. Those genes whose protein coding domain region information was not available in the protein database were excluded from this line of research, as well as the genes' whose domain region spanned its entirety.

The genomic data provided by ensemble protists for *B. natans* had gene identifiers unique to the Joint Genome Institute (JGI) databases, and therefore, the identifiers for *B. natans* top 200 expressed genes were not able to be searched for within the NCBI's protein database. To overcome this, each of the 200 genes were placed separately into the online tool ExPASy (Gasteiger et al., 2003). ExPASy translates the genes' nucleotide sequence into all six potential protein sequences. Any open reading frames (ORFs) were then highlighted by the program, these ORFs were then placed into the NCBI program Conserved Domain Search in order to check their validity (Marchler-Bauer et al., 2017). Valid ORFs were then searched for using a standard protein BLAST (Altschul et al., 1990), another NCBI program, this search revealed the gene within NCBI's protein database allowing for its protein coding domains to be located.

2.4 Statistical Analysis

The majority of statistical analysis performed on the codons of the chosen 33 organisms were carried out with the use of CodonW v. 1.4.2 (Peden, 1999). This included correspondence analysis and determining the Fop, Nc and GC3s values for codons found in a given set of genes.

Due to the use of different genetic codes any statistical analysis conducted on the codons for the organisms *S. salmonicida* and *T. thermophila* was completed using the nuclear code of Ciliophora as a base which was provided within CodonW. Statistical analysis on the codons for all other organisms used the universal genetic code, again provided for by CodonW.

All graphs and tables showed in this document were produced through Excel 2016, along with any other calculations, unless stated otherwise.

2.4.1 Identification of Optimal Codons

Optimal codons for all 33 species were determined using two methods: correspondence analysis through CodonW and by comparing the 5% top and bottom expressed genes. Correspondence analysis was performed by inputting the complete transcriptomes of each species into CodonW, the estimated optimal codons for the species was then generated in the form of a fop.coa file. CodonW then determined the F_{op} values for each organism with the use of their respective fop.coa files. Values for both Nc and GC3s were also calculated in CodonW.

To determine optimal codons with the use of expression data the files containing the top and bottom 5% expressed genes for every organism were entered separately into CodonW. These files were used to determine the Nc and GC3s for the codons within them. The gene sequences ending in partial codons were then trimmed. The modified files were then placed back into CodonW and the genes were concatenated. This allowed for a .blk file to be produced containing a table showing the concatenated genes use of codons.

A two-tailed Fisher's exact test was then conducted using a 2-by-2 contingency table to calculate the significant differences, in the form of p-values, between the use of codons found in the top and bottom 5% expressed genes for each species. This was carried out using QuickCalcs, software provided by GraphPad (GraphPad Software, 2020). Any difference in codon usage which exhibited a p-value equal to or less than 0.001 was deemed as significant. Any codon which was found at a significantly higher rate in the top 5% when compared with the bottom 5% of expressed genes was deemed as optimal.

2.4.2 Codon Usage in Domains and Non-Domains

CodonW was also used to determine codon usage in both the domains and the non-domains of the non-excluded genes from the top 200 expressed genes for all 33 species. This analysis was conducted again by determining GC3s and Nc values. The gene sequences ending in partial codons were then trimmed and concatenated.

The concatenated genes were then used as an input to obtain a .blk file showing codon usage. The usage of individual ADAT codons in both the domains and the non-domains of each organism was then placed in a 2-by-2 contingency table using the same statistical analysis stated in section 2.4.1.

2.4.3 Optimal Codon Usage in Domain and Non-Domain Codons

CodonW records the designated optimal codons for sets of genes into fop.coa files. Custom fop.coa files were created for each organism used in the study based on the codon usage data attained from their top and bottom 5% expressed genes. The custom fop.coa files were then loaded into CodonW in order to calculate the F_{op} values for the functional domains and non-domain regions of the non-excluded top 200 expressed genes for each species. The F_{op} values for the top and bottom 5% expressed genes for each organism as well as the F_{op} value for their transcriptome were also calculated using the custom fop.coa files.

2.4.4 Usage of ADAT preferred Codons

The usage of all ADAT preferred codons as opposed to their synonymous codon counterparts within the top and bottom 5% expressed genes and the functional domains and non-domains of the non-excluded top 200 expressed genes were calculated for all investigated species. For each organism the difference between the usage of ADAT preferred codons in the top and bottom 5% expressed genes was tested for significance with the use of a 2-by-2 contingency table, the software for which provided by VassarStats (VassarStats, 2020), as well as the significant difference between the functional domains and non-domains. In this case a Yates's chi-squared test was conducted by VassarStats due to the large size of the input values. A p-value equal to or less than 0.001 was considered to be significant.

2.5 Identification of tRNA Genes

The copy numbers for each type of tRNA within each of the 33 organisms were found using tRNAscan-SE, an online resource capable of searching through regions of DNA for tRNA genes (Lowe & Chan, 2016). Each organism's genomic file was entered into the tRNAscan-SE search box in FASTA format. However, due to tRNAscan-SE being limited to queries that were less than five million nucleotides, organism's genomic files had to be separated and searched in multiple subdivisions. These subdivisions were made up of roughly 4.5 million nucleotides each, depending on the size of super-contigs. The tRNA genes found for each organism from the individual searches were then compiled into one data set when all runs were complete.

In cases where individual super-contigs exceeded five million nucleotides, the contents of the contig was separated into two, with both halves of the contig being searched individually. A third search was then carried out which incorporated the middle of the whole super-contig, in case any tRNA genes were separated when the contig was split. The tRNA genes found from these three searches were then compared to ensure no tRNA genes were recorded twice.

2.6 Evidence of Deamination

In order to identify the presence of tRNA modified by the A-to-I deamination, small RNA molecules were screened from organisms with available data. These organisms were: *E. histolytica* (amoebozoan), *D. discoideum* (amoebozoan), *F. alba* (opisthokont), *C. owczarzaki* (opisthokont), *S. arctica* (opisthokont), *L. major* (excavate), *T. brucei* (excavate), *P. falciparum* (alveolate), *P. infestans* (stramenopile), *P. sojae* (stramenopile) and *T. pseudonana* (stramenopile). Any TAPSILVR tRNA genes found within the organisms' genetic data by tRNA-scan which possessed anticodons with adenosine at their wobble position were isolated from the genome.

These isolated genes were screened for within the small tRNA data, along with modified versions of the genes. The modified version of the gene included a guanosine at the 34th position replacing A, as inosine is replaced with guanosine during the production of cDNA. Any $\text{tRNA}^{\text{Ala}}_{\text{AGC}}$ genes that housed an A at the 37th position of the gene were modified three times, with each modified gene housing a different variation of potential post transcriptional modifications. The A at the 34th position was replaced with a G in the first modified gene (mod1), the A at the 37th position was replaced with a G in the second modified gene (mod2) and for the third modified gene both the A's at the 34th position and the 37th position were replaced with G's (mod3). This line of research was introduced in order to better understand the relationship between the ADAT enzymes: hetADAT and ADAT1.

SMALT was utilised to map the ncRNA or miRNA reads, dependent on the organism, onto each organisms collection of modified and unmodified genes using the same method as stated in section 2.2. Tablet was then used to observe the mapped reads, where the presence of both the modified and unmodified tRNA molecules was noted.

3.0 Results

Due to the number of species within this study, when represented within tables they have been separated, via their supergroups, into different coloured categories: red for the amoebozoans, orange for the apusozoan, yellow for the opisthokonts, green for the excavates, blue for the alveolates, purple for the stramenopiles and plum for the rhizarians.

3.1 The Degree of Codon Usage Bias in the Protists

With the use of Nc values, the degree of CUB was determined for all the genes present in the transcriptomes of the 33 protists investigated in this study. For each organism the average Nc value of their genes was calculated (Table 2), these averages ranged from 31.87 in the stramenopile *A. anophagefferens* to 56.31 in the excavate *T. brucei*; whereas the Nc values of individual genes within the transcriptomes of the organisms ranged from 20 to 61. Along with Nc values, GC3s values were calculated for the genes present within the transcriptomes. GC3s reveals the direction of bias, towards either AT-ending codons or towards GC-ending codons. The mean GC3s was calculated for all organisms (Table 2), ranging from 0.11 for the amoebozoan *E. histolytica* to 0.91 for the stramenopile *A. anophagefferens*. This depicts a strong bias towards AT ending codons in *E. histolytica* and an even stronger bias towards GC ending codons within *A. anophagefferens*. There were nine organisms whose average GC3s value was close to or precisely 0.5, including at least one organism from each supergroup. The majority of the opisthokonts, including *T. trahens*, and the stramenopiles have a high average GC3s value, whilst the majority of the alveolates show a low average GC3s value. There is a large variation in the excavates and amoebozoans’ average GC3s values, therefore, no trends can be seen.

Organism	Nc (±SD)	GC3s (±SD)	Number of genes
A. castellanii	42.99 ± 7.35	0.73 ± 0.15	14975
E. histolytica	33.82 ± 3.92	0.11 ± 0.05	8110
D. discoideum	33.00 ± 5.23	0.14 ± 0.08	13050
P. fungivorum	52.10 ± 5.43	0.54 ± 0.08	17125
T. trahens	38.71 ± 5.08	0.77 ± 0.07	10625
F. alba	38.24 ± 6.14	0.81 ± 0.09	6310
C. owczarzaki	47.85 ± 6.42	0.65 ± 0.08	10121
S. arctica	54.42 ± 6.35	0.54 ± 0.08	18731
G. lambila	53.58 ± 7.52	0.53 ± 0.10	7247
S. salmonicida	46.48 ± 7.86	0.37 ± 0.21	8320
L. major	44.94 ± 6.25	0.73 ± 0.08	8316
L. pyrrhocoris	44.60 ± 7.55	0.72 ± 0.13	13894
L. seymouri	47.91 ± 6.18	0.69 ± 0.09	8486
T. brucei	56.31 ± 4.71	0.50 ± 0.07	8728
T. cruzi	53.87 ± 5.23	0.54 ± 0.09	32274
T. theileri	50.36 ± 5.50	0.37 ± 0.13	11312
T. thermophila	41.60 ± 4.79	0.23 ± 0.06	24726
S. coeruleus	43.91 ± 4.33	0.23 ± 0.05	31427
P. gallinaceum	38.24 ± 5.77	0.18 ± 0.08	5247
P. falciparum	38.49 ± 4.21	0.16 ± 0.05	5354
P. vivax	51.33 ± 0.14	0.53 ± 0.14	5390
P. infestans	53.05 ± 5.49	0.54 ± 0.11	17764
P. sojae	46.32 ± 9.38	0.70 ± 0.14	26459
P. halstedii	54.08 ± 5.51	0.44 ± 0.07	15470
A. astaci	46.99 ± 5.35	0.71 ± 0.06	14015
A. invadans	47.51 ± 5.23	0.69 ± 0.06	10900
S. parasitica	37.96 ± 5.55	0.80 ± 0.07	20122
P. multistriata	51.13 ± 7.67	0.59 ± 0.15	12040
T. pseudonana	54.10 ± 7.80	0.47 ± 0.06	11673
N. gaditana	52.85 ± 6.11	0.61 ± 0.12	10930
A. anophagefferens	31.87 ± 9.51	0.91 ± 0.16	11521
B. natans	55.99 ± 4.88	0.50 ± 0.09	21685
P. brassicae	45.31 ± 7.80	0.73 ± 0.11	9728

Table 2 | Average Codon Usage Statistics in the Transcriptomes of the 33 Protists. All values to two decimal places. See Figures 5 to 37 in the appendix for a visualisation of the data shown within the table. Note – Nc, the effective number of codons; GC3s, guanine + cytosine content at synonymous third position of codon; SD, standard deviation.

Additionally, the average Nc and GC3s values were found separately for the top and bottom 5% expressed genes. These values were also found for the protein coding domains and non-domains of the non-excluded top 200 expressed genes for each species (Table 3).

If the CUB of an organism is influenced by the selection for the codons that provide translational efficiency, the highly expressed genes for that organism would show a greater bias, therefore, a lower Nc value than their weakly expressed counterparts. This can be seen in the vast majority of the protists where the top 5% expressed genes Nc value was lower than that of the bottom 5%. However, within the following protists, a lower Nc value was seen in the bottom 5% expressed genes: *D. discoideum* (amoebozoan), *L. seymouri* (excavate), *S. coeruleus* (alveolate), *P. multistriata* (stramenopile), *N. gaditana* (stramenopile) and *A. anophagefferens* (stramenopile) (Table 3).

Similarly, if codons which provide the greatest translational accuracy are selected for within an organism then the protein coding domain regions of genes will exhibit lower Nc values than their non-domain counterparts. The majority of protists show a stronger bias within their domain-encoding regions when compared to their non-domain regions whilst *E. histolytica* (amoebozoan), *T. brucei* (excavate), *T. cruzi* (excavate), *T. theileri* (excavate), *P. gallinaceum* (alveolate) and *P. vivax* (alveolate) show the opposite bias (Table 3).

	Top 5% Expressed Genes		Bottom 5% Expressed Genes		Protein Coding Domains		Non-Protein Coding Domains	
Organism	**Nc**	**GC3s**	**Nc**	**GC3s**	**Nc**	**GC3s**	**Nc**	**GC3s**
A. castellanii	34.66	0.85	39.51	0.81	31.33	0.87	33.74	0.81
E. histolytica	31.24	0.17	33.03	0.10	30.64	0.19	29.52	0.24
D. discoideum	32.46	0.21	31.02	0.10	31.84	0.28	32.26	0.25
P. fungivorum	52.25	0.55	53.22	0.52	50.66	0.58	51.71	0.54
T. trahens	34.94	0.82	40.48	0.75	31.96	0.84	35.46	0.80
F. alba	30.34	0.83	38.23	0.83	27.49	0.82	30.94	0.77
C. owczarzaki	36.90	0.75	50.70	0.61	32.85	0.79	36.26	0.71
S. arctica	51.50	0.55	53.95	0.50	44.89	0.58	46.96	0.53
G. lambila	50.63	0.60	57.02	0.50	48.99	0.65	50.43	0.62
S. salmonicida	41.45	0.52	45.28	0.24	36.93	0.59	36.98	0.55
L. major	35.71	0.82	46.11	0.72	35.99	0.82	39.49	0.76
L. pyrrhocoris	33.44	0.84	44.88	0.73	31.26	0.88	36.39	0.80
L. seymouri	47.67	0.70	42.93	0.74	44.40	0.72	47.15	0.69
T. brucei	51.89	0.52	57.33	0.49	50.88	0.57	50.50	0.54
T. cruzi	50.27	0.59	52.64	0.59	46.48	0.64	46.25	0.59
T. theileri	46.74	0.57	50.13	0.36	45.82	0.60	42.80	0.56
T. thermophila	36.01	0.31	39.78	0.21	31.63	0.42	34.29	0.37
S. coeruleus	44.19	0.26	42.39	0.21	41.11	0.27	41.16	0.29
P. gallinaceum	34.17	0.11	36.86	0.15	33.83	0.12	32.65	0.12
P. falciparum	37.21	0.17	39.94	0.17	38.00	0.15	38.26	0.17
P. vivax	49.91	0.55	52.94	0.54	48.50	0.54	46.65	0.55
P. infestans	41.42	0.74	54.41	0.56	35.22	0.79	35.46	0.77
P. sojae	32.24	0.88	49.91	0.66	28.88	0.92	30.74	0.86
P. halstedii	52.05	0.49	54.08	0.44	47.22	0.53	47.68	0.51
A. astaci	42.36	0.74	49.99	0.66	-	-	-	-
A. invadans	42.37	0.72	50.51	0.65	38.53	0.75	39.72	0.70
S. parasitica	31.91	0.85	37.63	0.81	28.04	0.85	31.46	0.78
P. multistriata	46.53	0.57	45.76	0.69	39.80	0.59	39.98	0.58
T. pseudonana	50.82	0.47	54.55	0.46	47.99	0.49	48.54	0.49
N. gaditana	52.40	0.58	51.21	0.63	46.83	0.69	49.74	0.59
A. anophagefferens	33.20	0.86	31.92	0.91	31.65	0.89	33.66	0.87
B. natans	52.14	0.50	54.49	0.47	45.25	0.49	49.25	0.35
P. brassicae	39.04	0.80	46.31	0.73	35.03	0.84	35.99	0.80

Table 3 | Nc and GC3s mean values for the top and bottom 5% expressed genes and the protein coding domains and non-domains for the 33 protists. . See Figures 5 to 37 in the appendix for a visualisation of the data shown within the table. Note – Nc, the effective number of codons; GC3s, guanine + cytosine content at synonymous third position of codon.

To visualise the degree of CUB within the transcriptomes of the 33 organisms Nc plots were produced (Appendix, Figures 5 - 37). Figure 4 shows genes present in the transcriptomes of *E. histolytica*, *T. brucei* and *T. trahens*, all displaying different degrees of codon bias.

The genes within the amoebozoan *E. histolytica* are shown to have a strong bias towards AT-ending codons, with highly biased genes showing lower GC3s values (Figure 3, A). This pattern can also be seen in *D. discoideum* (Figure 7) and in the majority of the researched alveolates (Appendix, Figures 21-24), with the exception of *P. vivax* (Appendix, Figure 25). The opposite can be seen in the transcriptome of *T. trahens* (apusozoan) where the genes showed a strong bias towards GC-ending codons, with highly biased genes featuring higher GC3s values (Figure 3, C). This pattern can also be seen in: *A. castellanii* (amoebozoan), *F. alba* (opisthokont), *C. owczarzaki* (opisthokont), *L. major* (excavate), *L. pyrrhocoris* (excavate), *L. seymouri* (excavate), *P. sojae* (stramenopile), *A. astaci* (stramenopile), *A. invadans* (stramenopile), *S. parasitica* (stramenopile), *P. multistriata* (stramenopile), *A. anophagefferens* (stramenopile) and *P. brassicae* (rhizarian) (Appendix, Figures 5, 10, 11, 15 - 17, 27, 29 – 32, 35 and 37). On the contrary, the genes within the excavate *T. brucei* showed very little indication of a relationship between GC3s and Nc values (Figure 3, B), along with *P. fungivorum* (amoebozoan), *S. arctica* (Opisthokont), *G. lamblia* (excavate), *P. halstedii* (stramenopile), *T. pseudonana* (stramenopile) and *B. natans* (rhizarian) (Appendix, Figures 8, 12, 13, 28, 33 and 36).

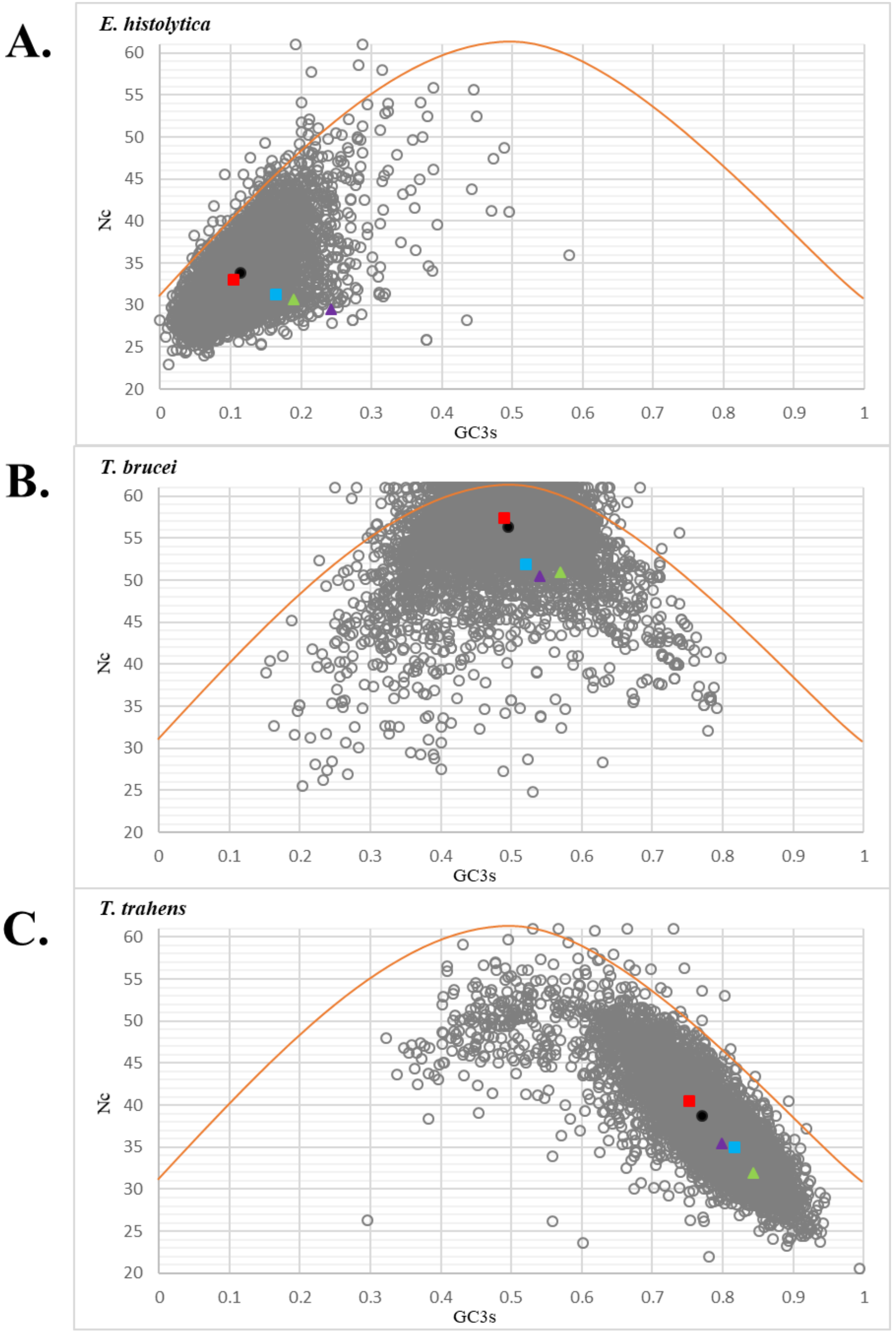
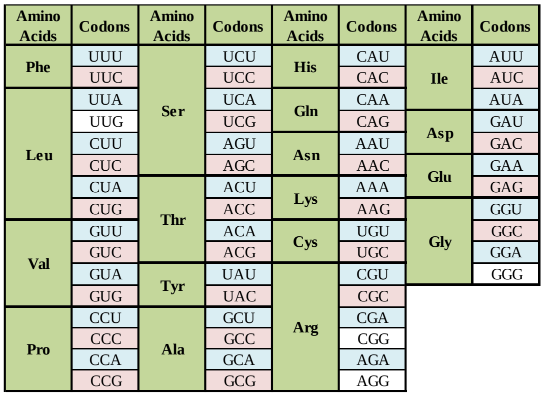


Figure 4 | Graphs showing genome wide Nc plots for *E. histolytica, T. brucei* and *T. trahens*. The Nc and GC3s values of all genes found in the transcriptome for each organism are shown as grey hollow circles. The average Nc and GC3s values for each organisms' set of genes are represented by a black dot. The average Nc and GC3s values for the top and bottom 5% expressed genes for each organism are represented by blue and red squares respectively. The average Nc and GC3s values for the found protein coding domains and non-domains of the top 200 expressed genes for each organism are represented by green and purple triangles respectively. The red bell-shaped curve represents the expected position of genes if their CUB was solely influenced by a neutral mutational bias (Wright, 1990). These particular Nc plots were chosen to illustrate the three different variations of patterns observed. The Nc plots for all genomes can be found in the appendix (Figures 5 to 37).

3.2 Optimal Codon Identification

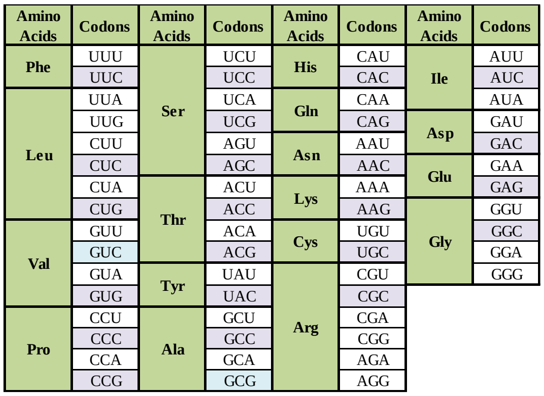
The optimal codons for each of the 33 protists were determined by performing correspondence analysis with the use of CodonW (Appendix, Tables 12–42). Optimal codons were also identified by carrying out a comparison between the codon usage of the top and bottom 5% expressed genes for each organism (Appendix, Tables 43-75). The two methods used identified differing optimal codons for all of the protists examined although the size of this difference varied greatly (Appendix, Tables 108-140). Whilst the two methods for finding optimal codons in *P. sojae* (stramenopile) only differed by 2 codons (Tables 4B) and those found for *T. theileri* (excavate)(Table 4A) differed completely.

A.



Amino Acids	Codons	Amino Acids	Codons	Amino Acids	Codons	Amino Acids	Codons
Phe	🟦 UUU	Ser	🟦 UCU	His	🟦 CAU	Ile	🟦 AUU
	🟥 UUC		🟥 UCC		🟥 CAC		🟥 AUC
Leu	🟦 UUA		🟦 UCA	Gln	🟦 CAA		🟦 AUA
	UUG		🟥 UCG		🟥 CAG	Asp	🟦 GAU
	🟦 CUU		🟦 AGU	Asn	🟦 AAU		🟥 GAC
	🟥 CUC		🟥 AGC		🟥 AAC	Glu	🟦 GAA
	🟦 CUA	Thr	🟦 ACU	Lys	🟦 AAA		🟥 GAG
	🟥 CUG		🟥 ACC		🟥 AAG	Gly	🟦 GGU
Val	🟦 GUU		🟦 ACA	Cys	🟦 UGU		🟥 GGC
	🟥 GUC		🟥 ACG		🟥 UGC		🟦 GGA
	🟦 GUA	Tyr	🟦 UAU	Arg	🟦 CGU		GGG
	🟥 GUG		🟥 UAC		🟥 CGC		
Pro	🟦 CCU	Ala	🟦 GCU		🟦 CGA		
	🟥 CCC		🟥 GCC		CGG		
	🟦 CCA		🟦 GCA		🟦 AGA		
	🟥 CCG		🟥 GCG		AGG		

B.



Amino Acids	Codons	Amino Acids	Codons	Amino Acids	Codons	Amino Acids	Codons
Phe	UUU	Ser	UCU	His	CAU	Ile	AUU
	🟪 UUC		🟪 UCC		🟪 CAC		🟪 AUC
Leu	UUA		UCA	Gln	CAA		AUA
	UUG		🟪 UCG		🟪 CAG	Asp	GAU
	CUU		AGU	Asn	AAU		🟪 GAC
	🟪 CUC		🟪 AGC		🟪 AAC	Glu	GAA
	CUA	Thr	ACU	Lys	AAA		🟪 GAG
	🟪 CUG		🟪 ACC		🟪 AAG	Gly	GGU
Val	GUU		ACA	Cys	UGU		🟪 GGC
	🟦 GUC		🟪 ACG		🟪 UGC		GGA
	GUA	Tyr	UAU	Arg	CGU		GGG
	🟪 GUG		🟪 UAC		🟪 CGC		
Pro	CCU	Ala	GCU		CGA		
	🟪 CCC		🟪 GCC		CGG		
	CCA		GCA		AGA		
	🟪 CCG		🟦 GCG		AGG		

Table 4A and 4B | Optimal Codons found for the excavate *T. theileri* (A) and the stramenopile *P. sojae* (B) using two methods: Correspondence Analysis (blue) and Expression Level (red). Codons shown with a blue background are those determined optimal by CodonW with the use of a correspondence analysis. Codons shown with a red background are those found to be optimal when comparing the top and bottom 5% expressed genes. Codons shown with a purple background are those found to be optimal by both methods.

Due to the extreme differences in optimal codons found by correspondence analysis and by expression data in certain organisms, there must be some optimal codons that are false positives. As optimal codons identified by correspondence analysis are only considered genuine if highly biased genes are those that are highly expressed (Peden, 1999), a heavy influence of mutational pressure on the CUB of the organisms could invalidate this method. This can be seen within *T. theileri* as its CUB is heavily influenced by a mutational bias towards AT-ending codons and its highly expressed genes exhibit higher GC3s values than the highly biased genes (Appendix, Figure 20). As it is hard to determine if mutation has a strong influence over CUB only those optimal codons found by expression data will be discussed and taken forward in further research.

The Fop values were found for each organisms genome and for the found protein coding domains and non-domains of the top 200 expressed genes for each organism (Table 5).

Organism	CDS (±SD)	Top 5% Expressed Genes (±SD)	Bottom 5% Expressed Genes (±SD)	Protein-Coding Domains (±SD)	Non-Domains (±SD)
A. castellanii	0.59 ± 0.10	0.77 ± 0.08	0.63 ± 0.07	0.82 ± 0.07	0.79 ± 0.12
E. histolytica	0.47 ± 0.07	0.66 ± 0.09	0.42 ± 0.04	0.68 ± 0.09	0.72 ± 0.11
D. discoideum	0.28 ± 0.09	0.45 ± 0.14	0.45 ± 0.06	0.52 ± 0.12	0.54 ± 0.12
P. fungivorum	0.43 ± 0.06	0.44 ± 0.06	0.41 ± 0.06	0.46 ± 0.05	0.42 ± 0.06
T. trahens	0.64 ± 0.10	0.73 ± 0.11	0.60 ± 0.08	0.76 ± 0.11	0.72 ± 0.12
F. alba	0.58 ± 0.11	0.80 ± 0.14	0.52 ± 0.06	0.85 ± 0.13	0.80 ± 0.16
C. owczarzaki	0.50 ± 0.10	0.69 ± 0.13	0.47 ± 0.07	0.75 ± 0.12	0.67 ± 0.16
S. arctica	0.46 ± 0.07	0.57 ± 0.11	0.44 ± 0.07	0.66 ± 0.12	0.62 ± 0.17
G. lambila	0.50 ± 0.09	0.57 ± 0.12	0.48 ± 0.07	0.63 ± 0.12	0.59 ± 0.14
S. salmonicida	0.41 ± 0.18	0.62 ± 0.18	0.28 ± 0.10	0.69 ± 0.17	0.66 ± 0.18
L. major	0.62 ± 0.08	0.75 ± 0.11	0.60 ± 0.05	0.73 ± 0.11	0.72 ± 0.14
L. pyrrhocoris	0.64 ± 0.13	0.80 ± 0.15	0.65 ± 0.08	0.84 ± 0.08	0.79 ± 0.15
L. seymouri	0.29 ± 0.08	0.29 ± 0.07	0.24 ± 0.11	0.26 ± 0.07	0.29 ± 0.08
T. brucei	0.36 ± 0.06	0.41 ± 0.11	0.35 ± 0.04	0.45 ± 0.11	0.45 ± 0.13
T. cruzi	0.53 ± 0.07	0.53 ± 0.08	0.49 ± 0.10	0.51 ± 0.09	0.58 ± 0.11
T. theileri	0.33 ± 0.12	0.54 ± 0.17	0.31 ± 0.11	0.57 ± 0.17	0.53 ± 0.16
T. thermophila	0.45 ± 0.08	0.67 ± 0.13	0.45 ± 0.06	0.81 ± 0.12	0.74 ± 0.17
S. coeruleus	0.38 ± 0.05	0.45 ± 0.06	0.36 ± 0.04	0.50 ± 0.08	0.48 ± 0.08
P. gallinaceum	0.62 ± 0.10	0.67 ± 0.05	0.62 ± 0.06	0.65 ± 0.05	0.68 ± 0.09
P. falciparum	0.36 ± 0.05	0.41 ± 0.06	0.34 ± 0.04	0.37 ± 0.06	0.35 ± 0.06
P. vivax	0.33 ± 0.06	0.38 ± 0.07	0.33 ± 0.04	0.38 ± 0.08	0.37 ± 0.11
P. infestans	0.46 ± 0.10	0.69 ± 0.10	0.50 ± 0.09	0.82 ± 0.07	0.79 ± 0.12
P. sojae	0.61 ± 0.14	0.81 ± 0.12	0.56 ± 0.10	0.87 ± 0.08	0.81 ± 0.15
P. halstedii	0.41 ± 0.07	0.52 ± 0.08	0.41 ± 0.07	0.58 ± 0.10	0.52 ± 0.13
A. astaci	0.71 ± 0.07	0.75 ± 0.07	0.65 ± 0.07	-	-
A. invadans	0.60 ± 0.06	0.65 ± 0.08	0.54 ± 0.07	0.70 ± 0.08	0.63 ± 0.12
S. parasitica	0.63 ± 0.08	0.74 ± 0.12	0.64 ± 0.08	0.81 ± 0.16	0.75 ± 0.20
P. multistriata	0.42 ± 0.13	0.49 ± 0.10	0.32 ± 0.14	0.38 ± 0.08	0.37 ± 0.11
T. pseudonana	0.34 ± 0.06	0.37 ± 0.09	0.33 ± 0.04	0.39 ± 0.11	0.36 ± 0.13
N. gaditana	0.40 ± 0.08	0.44 ± 0.09	0.38 ± 0.09	0.40 ± 0.09	0.45 ± 0.10
A. anophagefferens	0.22 ± 0.14	0.26 ± 0.18	0.23 ± 0.14	0.23 ± 0.17	0.27 ± 0.17
B. natans	0.46 ± 0.06	0.54 ± 0.10	0.45 ± 0.07	0.64 ± 0.15	0.36 ± 0.07
P. brassicae	0.63 ± 0.10	0.73 ± 0.09	0.62 ± 0.09	0.78 ± 0.09	0.74 ± 0.11

Table 5 | Frequencies of Optimal Codons (Fops) of the top and bottom 5% of expressed genes and the protein coding domains and non-domains of the top 200 expressed genes for each organism. Note – CDS, Coding Sequence; SD, standard deviation.

3.3 Codon Usage within the Protists using Expression Data

The protein coding domains and non-domains of the top 200 expressed genes were found for each organism. Due to reliance on the NCBI database not all domains could be found and therefore they could not be examined (Table 6). This was the case for the stramenopile *A. astaci* were no domains could be found due to there being no data available.

Organism	N° of D and ND
A. castellanii	105
E. histolytica	60
D. discoideum	56
P. fungivorum	105
T. trahens	111
F. alba	105
C. owczarzaki	122
S. arctica	74
G. lambila	93
S. salmonicida	90
L. major	61
L. pyrrhocoris	31
L. seymouri	99
T. brucei	66
T. cruzi	75
T. theileri	53
T. thermophila	70
S. coeruleus	85
P. gallinaceum	86
P. falciparum	115
P. vivax	98
P. infestans	81
P. sojae	88
P. halstedii	87
A. astaci	0
A. invadans	100
S. parasitica	70
P. multistriata	58
T. pseudonana	85
N. gaditana	81
A. anophagefferens	64
B. natans	57
P. brassicae	93

Table 6 | The number of Protein Coding Domains and Non-Domains for the top 200 expressed genes for the 33 protists. Note- N°, number; D, Protein-coding domains; ND, Non-domains.

The overall usage of the ADAT preferred codons was found for each organisms’ top and bottom 5% expressed genes (Table 7). If found at a higher frequency within the highly expressed genes then it can be concluded that collectively they are selected for due to translational efficiency. Although a p-value of 0.01 is commonly used in codon usage research as a cut off point for significant difference, within this study a value of 0.001 was chosen to improve upon the accuracy of the findings.

Organism	Top 5%	Bottom 5%	p-value
A. castellanii	0.56 (72440)	0.44 (72009)	p<0.0001
E. histolytica	0.08 (2406)	0.04 (1876)	p<0.0001
D. discoideum	0.18 (16104)	0.05 (7459)	p<0.0001
P. fungivorum	0.32 (73849)	0.29 (49801)	p<0.0001
T. trahens	0.53 (34081)	0.45 (69537)	p<0.0001
F. alba	0.62 (31208)	0.44 (67092)	p<0.0001
C. owczarzaki	0.48 (35986)	0.30 (48547)	p<0.0001
S. arctica	0.27 (27897)	0.21 (19039)	p<0.0001
G. lambila	0.29 (24190)	0.22 (23878)	p<0.0001
S. salmonicida	0.29 (12951)	0.08 (3290)	p<0.0001
L. major	0.37 (22641)	0.27 (63129)	p<0.0001
L. pyrrhocoris	0.39 (36233)	0.31 (63939)	p<0.0001
L. seymouri	0.28 (44649)	0.31 (18051)	p<0.0001
T. brucei	0.24 (8608)	0.18 (30996)	p<0.0001
T. cruzi	0.23 (54426)	0.21 (54963)	p<0.0001
T. theileri	0.26 (19327)	0.13 (16200)	p<0.0001
T. thermophila	0.18 (35711)	0.08 (22185)	p<0.0001
S. coeruleus	0.11 (24634)	0.09 (23476)	p<0.0001
P. gallinaceum	0.04 (1531)	0.05 (2385)	p<0.0001
P. falciparum	0.08 (2469)	0.07 (2474)	p<0.0001
P. vivax	0.24 (10497)	0.24 (17719)	p=0.3594
P. infestans	0.33 (40078)	0.24 (35639)	p=0.0001
P. sojae	0.45 (62852)	0.29 (62565)	p<0.0001
P. halstedii	0.19 (17371)	0.18 (9264)	p<0.0001
A. astaci	0.39 (37914)	0.33 (35189)	p<0.0001
A. invadans	0.37 (30238)	0.32 (24883)	p<0.0001
S. parasitica	0.47 (59833)	0.39 (91918)	p<0.0001
P. multistriata	0.37 (23580)	0.40 (59399)	p<0.0001
T. pseudonana	0.25 (14735)	0.21 (35231)	p<0.0001
N. gaditana	0.29 (15050)	0.32 (27715)	p<0.0001
A. anophagefferens	0.48 (16989)	0.46 (62910)	p<0.0001
B. natans	0.23 (37565)	0.19 (38824)	p<0.0001
P. brassicae	0.39 (27016)	0.35 (36806)	p<0.0001

Table 7 | The use of ADAT preferred codons within the top and bottom 5% of expressed genes for the 33 protists. The codon usage data is given in two forms, inside and outside of brackets. The numbers outside the brackets represent the proportion of ADAT-preferred codons found within the set of genes as opposed to synonymous codons. The numbers inside of the brackets represent the number of ADAT-preferred codons found within the set of genes. The p-values shown were calculated by a 2x2 contingency table. Any p-value above 0.001 was deemed insignificant, these are shown using red text.

The overall usage of the ADAT preferred codons was also found for each organisms' protein coding domains and non-domains (Table 8). This was carried out to determine if they are collectively selected for due to translational accuracy, if this was the case they would be found at a greater frequency within their protein coding domains. Again, data which had a p-value above 0.001 were deemed not significantly different to improve accuracy.

Organism	Domains	Non-Domains	p-value
A. castellanii	0.61 (7400)	0.55 (2442)	p<0.0001
E. histolytica	0.09 (228)	0.14 (166)	p<0.0001
D. discoideum	0.26 (1000)	0.17 (580)	p<0.0001
P. fungivorum	0.36 (7690)	0.30 (5178)	p<0.0001
T. trahens	0.59 (4161)	0.47 (2563)	p<0.0001
F. alba	0.69 (7681)	0.57 (2905)	p<0.0001
C. owczarzaki	0.58 (5367)	0.43 (2373)	p<0.0001
S. arctica	0.30 (1744)	0.26 (545)	p<0.0001
G. lambila	0.35 (6082)	0.28 (3230)	p<0.0001
S. salmonicida	0.42 (2579)	0.33 (728)	p<0.0001
L. major	0.39 (1715)	0.32 (910)	p<0.0001
L. pyrrhocoris	0.40 (1300)	0.35 (515)	p=0.0027
L. seymouri	0.30 (4140)	0.27 (6364)	p<0.0001
T. brucei	0.31 (1427)	0.28 (656)	p=0.0042
T. cruzi	0.25 (2163)	0.23 (880)	p=0.0361
T. theileri	0.27 (1421)	0.24 (479)	p=0.0019
T. thermophila	0.30 (2569)	0.24 (763)	p<0.0001
S. coeruleus	0.14 (1058)	0.12 (346)	p=0.0100
P. gallinaceum	0.04 (307)	0.08 (253)	p<0.0001
P. falciparum	0.07 (689)	0.06 (759)	p=0.1703
P. vivax	0.23 (2315)	0.23 (1517)	p=0.8875
P. infestans	0.36 (2785)	0.37 (1076)	p=0.6468
P. sojae	0.53 (2968)	0.50 (1232)	p=0.0029
P. halstedii	0.21 (1538)	0.20 (710)	p=0.3929
A. invadans	0.40 (3431)	0.39 (2096)	p=0.0595
S. parasitica	0.56 (3041)	0.43 (926)	p<0.0001
P. multistriata	0.44 (2591)	0.42 (951)	p=0.2579
T. pseudonana	0.28 (1617)	0.26 (1228)	p=0.0171
N. gaditana	0.37 (2226)	0.28 (934)	p<0.0001
A. anophagefferens	0.49 (2002)	0.44 (2399)	p<0.0001
B. natans	0.25 (2115)	0.15 (262)	p<0.0001
P. brassicae	0.44 (3914)	0.40 (1093)	p=0.0003

Table 8 | The use of ADAT preferred codons within the found protein coding domains and non-domains of the top 200 expressed genes for 32 protists with the exclusion of *A. astaci*. The codon usage data is given in two forms, inside and outside of brackets. The numbers inside of the brackets represent the number of ADAT-preferred codons found within the set of genes. The numbers outside the brackets represent the proportion of ADAT-preferred codons found within the set of genes as opposed to synonymous codons. The p-values shown were calculated by a 2x2 contingency table, any p-value above 0.001 was deemed insignificant, these are shown using red text. *A. astaci* was excluded from this table due to the lack of data available.

3.4 tRNA abundance

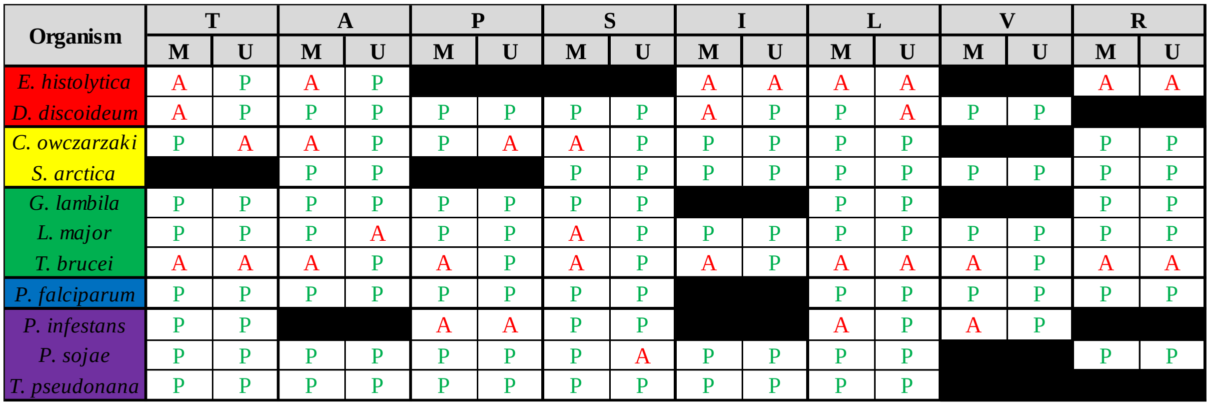
As previously mentioned the number of genes present within a genome for a particular tRNA type positively correlates to the tRNA molecule abundance (Iben & Maraia, 2012). Therefore, the tRNA with the highest copy number for each amino acid is classified as the major tRNA gene and is expected to be the most abundant tRNA molecule for that amino acid. Under a selection model, the optimal codons are those which correspond to the most abundant tRNA molecules.

tRNAscan-SE was used to identify the tRNA genes from each species. From these the most abundant/only tRNA for each amino acid in each species was acknowledged as a major tRNA gene. Due to having a wide range of organisms, the number of tRNA genes found for each species varied greatly; with 4471 tRNA genes found within the stramenopile *P. infestans* and only 41 tRNA genes found within the alveolate *P. gallinaceum*.

All tRNA genes found, including those which were chimeric and pseudo, were initially recorded (Appendix, Tables 142-174). However, chimeric and pseudo tRNA genes were removed from further study in order to maximise accuracy (Appendix, Tables 175-207. In some organisms like the excavate *T. brucei* and the alveolate *P. falciparum* the removal of these tRNA genes was minimal, with only 2.3% and 3.2% of their tRNA genes being removed respectively. However, in the three stramenopiles *P. halstedii, A. astaci* and *N. gaditana* over 50% of their genes were removed. The consequence of the removal of chimeric and pseudo tRNA genes was a change in 19% of the major tRNA genes. The identified tRNA genes were then used to compare the major tRNA genes with the optimal codons found within each species. A comparison could not be made when no tRNA genes were found for an amino acid or if no optimal codon could be identified. Additionally, if all tRNA genes for an amino acid were found at an equal frequency, meaning there were no major tRNA genes, a comparison could not be made.

3.5 The Presence of the A-to-I Deamination

The modified and unmodified versions of TAPSILVR tRNA molecules with adenosine at the wobble position were screened for within nc/miRNA reads found from NCBI for 11 of the protists (Table 9). There were some TAPSILVR tRNA molecules with adenosine at the wobble position that could not be screened due to the absence of the A34 tRNA genes within certain organisms. Although alanine, valine and arginine A34 tRNA molecules were present within *P. infestans* the isoaccepting G34 tRNAs were also found to be present, hence, the screening for the modified tRNA molecules for these amino acids could be carried out.



Organism	**T**		**A**		**P**		**S**		**I**		**L**		**V**		**R**	
	M	**U**	**M**	**U**	**M**	**U**	**M**	**U**	**M**	**U**	**M**	**U**	**M**	**U**	**M**	**U**
E. histolytica	A	P	A	P	■	■	■	■	A	A	A	A	■	■	A	A
D. discoideum	A	P	P	P	P	P	P	P	A	P	P	A	P	P	■	■
C. owczarzaki	P	A	A	P	P	A	A	P	P	P	P	P	■	■	P	P
S. arctica	■	■	P	P	■	■	P	P	P	P	P	P	P	P	P	P
G. lambila	P	P	P	P	P	P	P	P	■	■	P	P	■	■	P	P
L. major	P	P	P	A	P	P	A	P	P	P	P	P	P	P	P	P
T. brucei	A	A	A	P	A	P	A	P	A	P	A	A	A	P	A	A
P. falciparum	P	P	P	P	P	P	P	P	■	■	P	P	P	P	P	P
P. infestans	P	P	■	■	A	A	P	P	■	■	A	P	A	P	■	■
P. sojae	P	P	P	P	P	P	P	A	P	P	P	P	■	■	P	P
T. pseudonana	P	P	P	P	P	P	P	P	P	P	P	P	■	■	■	■

Table 9 | The Presence or Absence of the A34-to-I34 modified and unmodified TAPSILVR tRNA molecules in 11 of the Protists. Black spaces are used to show a lack of data availability for screening. Note – M, modified TAPSILVR tRNA molecule; U, unmodified TAPSILVR tRNA molecule; A, absence of the tRNA molecule (red); P, presence of the tRNA molecule (green).

The presence of the A37-to-I37 modification was also screened for using modified versions of tRNA$^{\text{Ala}}_{\text{AGC}}$ molecules from nine organisms (Table 10). There were no A37 tRNA$^{\text{Ala}}_{\text{AGC}}$ molecules in the alveolate *P. falciparum* and the stramenopile *P. infestans*.

Organism	Modifications	
	M37	**M34 + M37**
E. histolytica	A	A
D. discoideum	A	P
C. owczarzaki	A	A
S. arctica	A	A
G. lambila	A	A
L. major	A	A
T. brucei	A	A
P. sojae	A	P
T. pseudonana	A	P

Table 10 | The Presence or Absence of the A37-to-I37 Modification in relation to the Presence of the A34-to-I34 Modification in Nine Protists. Note – M37, A37-to-I37 modified Alanine tRNA molecule; M34 + M37, Alanine tRNA molecule with both the A34-to-I34 modification and the A37-to-I37 modification; A, absence of the modified tRNA molecule (red); P, presence of the tRNA molecule (green).

4.0 Discussion

The study was designed to provide an accurate estimation into the origins of the TAPSILVR I34 tRNA deamination in eukaryotes by looking at its distribution across the eukaryotes. This was achieved by performing various bioinformatical analyses on the genomes of 33 protists. Using the results of these techniques, statistical analysis allowed for the calculation for the usage of ADAT preferred codons within the genomes of the 33 protists. The usage of ADAT preferred codons was then compared with the tRNA genes for each protist; using this comparison the occurrence of the TAPSILVR I34 tRNA deamination within each protist could be determined. This data was used in conjunction with the presence of tRNA modified by the A-to-I deamination to accurately estimate the origin of the deamination in eukaryotes.

The usage of the ADAT preferred codons within different sections of the genomes of each protist was also calculated. The usage of the ADAT preferred codons within the protists' 5% highly expressed genes was compared to that within their 5% weakly expressed genes, in order to determine if the codons were selected for due to translational efficiency of the codons. While a comparison was made between the usage of the ADAT preferred codons within the gene regions that encode the functional domains and that encode the non-domain amino acids of the top 200 highly expressed genes of each protist to determine if translational accuracy of the codons provided a selectional advantage. Therefore, where data were available, the mechanisms behind the selection of the ADAT preferred codons could be estimated.

As a tertiary aim this study hoped to better understand the relationship between ADAT1 and hetADAT enzymes and their respective modifications. This was carried out by searching for the presence of three different modifications of the gene for $tRNA^{Ala}_{AGC}$. As $tRNA^{Ala}_{AGC}$ genes could only be identified in nine out of the 33 organisms, more data is needed to fully understand the relationship between the ADAT1 and hetADAT enzymes.

A third ADAT enzyme has also been found in eukaryotes, ADAT1. Instead of A34, ADAT1 modifies A37 into I37 in $tRNA^{Ala}$ only. It has recently been suggested in Torres (2014) that the modification of A37 by ADAT1 is independent of the presence of I34. The evidence does suggest that the I34 modification can and does occur independently of the I37 modification (Maas, Gerber & Rich, 1999), although the supporting literature provided in the review does not provide evidence that the I37 modification is independent of the I34 modification. Therefore, it is unclear if ADAT1's interaction depends on the interaction of hetADAT. ADAT1's interaction could also depend on the presence of I34, thereby only modifying the products of hetADAT activity.

4.1 The Overall CUB of the Protists in Relation to their Optimal Codons

The majority of optimal codons found for *A. castellanii* (amoebozoan), *T. trahens* (apusomonadan), *F. alba* (opisthokont), *C. owczarzaki* (opisthokont), *G. lamblia* (excavate), *L. major* (excavate), *L. pyrrhocoris* (excavate), *P. infestans* (stramenopile), *P. sojae* (stramenopile), *A. astaci* (stramenopile), *A. invadans* (stramenopile), *S. parasitica* (stramenopile), *B. natans* (rhizarian) and *P. brassicae* (rhizarian) end in either G or C which is consistent with the overall bias of their genomes (Appendix, Figures 5, 9-11, 13, 15, 16, 26, 27, 29-31, 36 and 37). This shows that the overall CUB of these organisms is directly influenced by selection for GC-ending codons.

The optimal codons found for *E. histolytica* (amoebozoan), *D. discoideum* (amoebozoan), *P. fungivorum* (amoebozoan), *S. arctica* (opisthokont), *S. salmonicida* (excavate), *T. brucei* (excavate), *T. theileri* (excavate), *T. thermophila* (alveolate), *S. coeruleus* (alveolate), *P. halstedii* (stramenopile) and *T. pseudonana* (stramenopile) are also majority GC-ending, however, this bias is not observed within their genomes (Appendix, Figures 6-8, 12, 14, 18, 20-22, 28 and 33). Therefore, the CUB of these organisms appears to be influenced by a mutational bias towards AT-ending codons whilst also being subjected to selection for GC-ending codons.

The majority of optimal codons found in *L. seymouri* (excavate), *T. cruzi* (excavate), *P. falciparum* (alveolate), *P. multistriata* (stramenopile), *N. gaditana* (stramenopile) and *A. anophagefferens* (stramenopile) are AT ending, which is not consistent with the GC3s values found within their genomes (Appendix, Figures 17, 19, 24, 32, 24 and 35). The CUB of these organisms must be influenced towards GC-ending codons via mutation, whilst selectional forces increase the use of AT-ending codons.

P. gallinaceum is the only protist to show a mutational bias towards AT-ending codons alongside and selectional bias towards AT-ending codons acting on its CUB (Appendix, Figure 23).

4.2 Fop Values found for the Protein Coding Domains and the Non-Domains

The Fop values found for the protein coding domains and non-domains revealed that within *E. histolytica* (amoebozoan), *D. discoideum* (amoebozoan), *L. seymouri* (excavate), *T. cruzi* (excavate), *P. gallinaceum* (alveolate), *N. gaditana* (stramenopile) and *A. anophagefferens* (stramenopile) optimal codons were found at a significantly higher frequency within their non-domains than their protein coding domains (Table 5). As optimal codons that increase translational accuracy would be found at a higher rate within the protein coding domains, evidence for selection due to translational accuracy could not be found within these organisms. Selection for translational accuracy also does not appear to affect the excavate *T. brucei* as there was no difference between the Fop of its protein coding domains and the Fop of its non-domains.

The other 24 organisms showed a higher Fop within their domains than their non-domains (Table 5). Therefore, within these protists their optimal codons appear to be selected for their translational accuracy. The significant increase between the Fops of domains and non-domains was subtle in the majority of these organisms, with a range from 0.01 to 0.08, with the exception of the rhizarian *B. natans*.

Within *B. natans* the Fop of its domain regions was higher by 0.28 than its non-domain regions. This shows that the selection for codons that provide increased translational accuracy is much stronger within *B. natans* than in any of the other protists studied.

4.3 Frequency of ADAT Preferred Codons within the Highly and Weakly Expressed Genes

In 28 of the 33 organisms the ADAT preferred codons were used at a significantly higher rate in their top 5% expressed genes as opposed to their bottom 5% expressed genes (Table 7). Therefore, the ADAT preferred codons are collectively selected for within these organisms to increase their translational efficiency.

Collectively, the ADAT preferred codons were used at a significantly higher rate in their bottom 5% expressed genes as opposed to their top 5% expressed genes within *L. seymouri* (excavate), *P. gallinaceum* (alveolate), *P. multistriata* (stramenopile) and *N. gaditana* (stramenopile) (Table 7). No significant difference was found when comparing the use of ADAT preferred codons in the top and bottom 5% expressed genes for *P. vivax*. Therefore, the ADAT preferred codons do not appear to provide a selectional advantage in terms of translational efficiency within these five protists.

4.4 Frequency of ADAT Preferred Codons within the Protein Coding Domains and the Non-Domains

Collectively the ADAT preferred codons were found at a significantly higher frequency in the domains of 17 of the 33 organisms when compared to the frequency within in their non-domains (Table 8). Therefore, the ADAT-preferred codons appear to be selected for within these organisms due to translational accuracy.

The ADAT preferred codons were collectively found at a significantly lower rate in the domains of the amoebozoan *E. histolytica* and the alveolate *P. gallinaceum* than their non-domains. This corresponds with the strong bias towards AT-ending codons acting upon the CUB of *P. gallinaceum* (Appendix, Figure 23). The lack of selection due to translational accuracy for the ADAT preferred codons within *E. histolytica* is not consistent with its optimal codons as they were GC-ending. There was no significant difference between the usage of ADAT preferred codons within the domains and non-domains of 13 of the protists. Therefore, showing that collectively, the ADAT preferred codons within these organisms were not selected for due to translational accuracy. However, this is assuming that the assumption that codons which are seen in higher frequencies within the protein coding domains as opposed to their non-domains is valid.

4.5 tRNA Genes

4.5.1 Chimeric tRNA

A significant number of tRNA genes found within the studied protist genomes were identified by tRNA-scan as chimeric. A tRNA gene is found to be a chimera when the anticodon found within the gene and the structure of the tRNA molecule correspond to different tRNA types (Perry, 2005).

The biological importance of these chimeric RNA molecules is not fully understood due to lack of research; however, they could serve to recode proteins given specific conditions (Lowe, 2016). At least one tRNA gene for tryptophan in 27 of the organisms was found to chimeric, in 17 of these cases all tryptophan tRNA genes were found to be either chimeric or pseudo (Appendix, Tables 142- 174). This calls into question the validity of tRNA-scan and how it identifies chimeric genes.

4.5.2 SeC tRNA

During the search for tRNA genes putative tRNA genes containing a TCA anticodon for the amino acid selenocysteine (tRNA^{SeC}) were found in 27 of the protist's genomes (Appendix, Table 141). In the majority of these species only a single tRNA^{SeC} gene was found, however, two were found in the stramenopile *T. pseudonana* whilst three were found in both the excavate *T. cruzi* and the stramenopile *P. multistriata*. As the gene for tRNA^{SeC} was found within at least one protist belonging to each of the supergroups looked at within this study, it can be concluded that unless an extreme case of convergent evolution took place, the tRNA^{SeC} gene is ancestral to the eukaryotes. Due to the anticodon of tRNA^{SeC} complementing the known STOP codon UGA, this codon may not indicate the end of genes in organisms where the tRNA^{SeC} gene was found. Therefore, this may suggest that genes within the genomes which possess SeC tRNA require re-evaluation and re-annotation.

4.5.3 Major tRNA Genes and the Optimal Codons for the Two-Fold Amino Acids

For the two-fold degenerate amino acids, 16 of the protists (*A. castellanii*, *E. histolytica*, *P. fungivorum*, *T. trahens*, *F. alba*, *C. owczarzaki*, *S. arctica*, *G. lamblia*, *L. major*, *L. pyrrhocoris*, *T. brucei*, *T. theileri*, *T. thermophila*, *P. vivax*, *P. infestans* and *P. sojae*) showed a perfect match between their optimal codons and their major tRNA genes. However, this was not the case for *L.*

seymouri (excavate), *P. gallinaceum* (alveolate) and *N. gaditana* (stramenopile) as the optimal codons for the two-fold amino acids did not match up with any of the major tRNA genes.

The A34 and G34 tRNA genes which correspond to the amino acids of glutamine, lysine and glutamic acid were present within the majority of the protists. With 30 out of the 33 protists having both the A34 and G34 tRNA genes for at least one of these amino acids. Further research into the occurrence of the isoaccepting tRNA genes is needed to understand the significance of these findings.

Glutamine in both the excavate *S. salmonicida* and the alveolate *T. thermophila* is a four-fold-degenerate amino acid. For both species the optimal codons for glutamine perfectly matched the apparent major tRNA gene. However, although the genetic code for tRNA isotype prediction can be changed in tRNAscan-SE to accommodate $tRNA_{TTA}$ and $tRNA_{CTA}$ this was not carried out, therefore, the most abundant tRNA molecule could not be determined for glutamine in *S. salmonicida* and *T. thermophila*.

4.5.4 Major tRNA Genes and the Optimal Codons for the TAPILVR Amino Acids

The codons for the TAPSILVR amino acids that were optimal in the highest number of organisms were the ADAT preferred codons (Appendix, Tables 208-213). However, the corresponding tRNA molecules were not seen to be the major tRNA molecule in any of the 33 species. Furthermore, the genes for the molecules $tRNA^{Thr}_{GGT}$, $tRNA^{Pro}_{GGG}$, $tRNA^{Ser}_{GGA}$ and $tRNA^{Leu}_{GAG}$ were absent in all of the 33 protists examined, whilst the genes for $tRNA^{Ala}_{GGC}$ and $tRNA^{Arg}_{GCG}$ were absent in nearly all the protists with the exception of the stramenopile *P. infestans*.

The major tRNAs for the TAPSILVR amino acids varied greatly throughout the organisms. The A34 tRNA gene for the TAPSILVR amino acids was found to be the major tRNA gene in the highest number of organisms compared to the other isoaccepting tRNA genes. This meant that for the TAPSILVR amino acids only 23% of their optimal codons matched perfectly with the major tRNA gene throughout the species. However, due to the high amount of ADAT preferred optimal codons found for the TAPSILVR amino acids this number increased to 49% when the A34-to-I34 tRNA deamination is considered.

4.6 Major tRNA genes compared to the optimal codons within the supergroups

4.6.1 Amoebozoa

The ADAT preferred codons for threonine, alanine, isoleucine and leucine were deemed optimal in all of the amoebozoans, despite the absence of genes for their complementary tRNA molecules. The ADAT preferred codons for serine and valine, with the exception of *E. histolytica*, and for proline with the exception of *D. discoideum*, were also found to be optimal despite having no complementary tRNA molecules throughout the amoebozoans. As the major tRNA genes were almost exclusively the A34 tRNA genes, this data is consistent with the presence of the I34 tRNA deamination (Table 11). When the I34 tRNA deamination is considered the amoebozoans optimal codons go from complementing the major tRNA 23% of the time to 60% of the time. Showing that the I34 tRNA deamination allows for the selection for the ADAT preferred codons within the Amoebozoa.

Furthermore, due to the ADAT preferred codons being found at higher frequencies within the highly expressed genes (HEG) than the weakly expressed genes (WEG) for the four amoebozoans their selection appears to be driven by translational efficiency (Table 7). The selection for the ADAT preferred codons within the Amoebozoa also appears to be driven by translational accuracy as, with the exception of *E. histolytica*, they were found at higher frequencies within their protein coding domains when compared to their non-domains (Table 8).

Amino Acid	*A. castellanti*	*E. histolytica*	*D. discoideum*	*P. fungivorum*
Thr	**ACU, ACC**	**ACU, ACC**	**ACU, ACC**	**ACC, ACG**
Ala	**GCU, GCC**	**GCU, GCC**	**GCU, GCC**	**GCC**
Pro	**CCC**	**CCA**	**CCA**	**CCC**
Ser	**UCC, UCG**	**UCC, UCA, GCU**	**UCU, UCC, GCU**	**UCC**
Ile	**AUC**	**AUU, AUC**	**AUC**	**AUC**
Leu	**CUU, CUC**	**CUU, CUC**	**CUU, CUC**	**CUC, CUG**
Val	**GUU, GUC**	**GUU, GUC**	**GUC**	**GUC**
Arg	**CGU, CGC, AGG**	**AGA**	**CGU**	**AGA**

Table 11 | Optimal Codons paired with the Major tRNA Genes for the Four Amoebozoans. Optimal codons that complement the major tRNA gene are written in blue. The optimal codons that complement the major tRNA gene if the adenosine at the wobble position underwent deamination to inosine are written in green.

4.6.2 Opisthokonts and T. trahens

The tRNA genes that complement the ADAT preferred codons were not found in any of the opisthokonts or within *T. trahens*. However, at least six of the ADAT preferred codons were found to be optimal within each of the species, with the ADAT preferred codons for threonine, alanine, leucine, valine and arginine being found as optimal in all of the opisthokonts and *T. trahens*. The lack of G34 tRNA genes for the TAPSILVR amino acids meant that only 25% of their optimal codons complemented the major tRNA gene. However, as the majority of the major tRNA genes were A34 tRNA, this percentage increases to 66% when the I34 tRNA deamination is considered (Table 12). Therefore, the presence of the I34 tRNA modification allows for the selection of the ADAT preferred codons within the opisthokonts and *T. trahens*.

Within both the opisthokonts and the apusozoan *T. trahens* the ADAT preferred codons were seen at significantly higher frequencies in the HEG of the respective organisms when compared to the WEG (Table 7). They were also seen at significantly higher frequencies in the protein coding domains of the organisms when compared to their non-domains (Table 8). Therefore, overall the ADAT preferred codons were shown to be selected for both translational efficiency and accuracy in the opisthokonts and *T. trahens*.

Amino Acid	*T. trahens*	*F. alba*	*C. owczarzaki*	*S. arctica*
Thr	**ACC, ACG**	**ACU, ACC**	**ACC**	**ACU, ACC**
Ala	**GCC**	**GCU, GCC**	**GCU, GCC**	**GCU, GCC**
Pro	**CCC, CCG**	**CCU, CCC**	**CCC**	**CCU, CCC**
Ser	**UCC, UCG**	**UCU, UCC**	**UCC, UCG**	**UCU, UCC**
Ile	**AUC**	**-**	**AUC**	**AUC**
Leu	**CUC**	**UUA, UUG, CUU, CUC**	**CUC**	**CUU, CUG**
Val	**GUC**	**GUU, GUC**	**GUC**	**GUC**
Arg	**CGU, CGC**	**CGU, CGC, AGA**	**CGC**	**CGU, CGC**

Table 12 | Optimal Codons paired with the Major tRNA Genes for the Apusomonadan *T. trahens* and Three Opisthokonts. Optimal codons that complement the major tRNA gene are written in blue. The optimal codons that complement the major tRNA gene if the adenosine at the wobble position underwent deamination to inosine are written in green.

4.6.3 Excavata

The genomes for both the excavates *G. lamblia* and *S. salmonicida* contain the gene for $tRNA^{Val}_{GAC}$. They were also the only excavates to have absence of the A34 tRNA gene for isoleucine. As both species are closely related, these similarities could be considered ancestral or a case of convergent evolution. All other G34 tRNA genes were absent from the excavates, whilst their complementary codons (ADAT preferred codons) were found to be optimal the majority of the time. In *L. seymouri*, however, none of the ADAT preferred codons were found to be optimal despite the protists' overall bias for GC-ending codons.

For the TAPSILVR amino acids within the excavates, 23% of their optimal codons complemented their major tRNA gene. However, consistent with the presence of the I34 deamination, this percentage increased to 48% when the A34 major tRNA genes could be paired with the ADAT preferred optimal codons (Table 13). Thus, the I34 deamination allows for the selection of the ADAT preferred codons within the majority of the excavates examined with the exception of *L. seymouri*.

In *L. seymouri* no ADAT preferred codons were found to be optimal. However, collectively they were found in significantly higher frequencies within the protein coding domains when compared to the non-domains (Table 8). This points towards the possibility that the optimal codons for translational efficiency are not the optimal codons for translational accuracy. This is supported with the significantly higher frequency of both ADAT preferred codons for alanine and arginine within the protein coding domains when compared to the non-domains (Appendix, Table 88). However, this line of thought is highly speculative as there have been no previous reports of an organism possessing different sets of optimal codons for translational efficiency and accuracy.

The excavates *G. lamblia, S. salmonicida* and *L. major* also show significantly higher frequencies of ADAT preferred codons within the protein coding domains when compared to the non-domains (Table 8). Thus, presenting evidence that the ADAT preferred codons are selected for translational accuracy. Translational efficiency also appeared to drive the selection for the ADAT preferred codons within these organisms, along with the other excavates, except *L. seymouri* due to the ADAT preferred codons being seen in significantly higher frequencies within their HEG than their WEG (Table 7).

Amino Acid	*G. lamblia*	*S. salmonicida*	*L. major*	*L. pyrrhocoris*
Thr	ACC, ACG	ACC, ACG	ACC, ACG	ACG
Ala	GCC, GCG	GCC, GCG	GCU	GCU, GCC, GCG
Pro	CCG	CCC	CCC, CCG	CCC, CCG
Ser	UCC, UCG, AGC	UCU, UCC, UCG, AGC	UCC	UCC, UCG
Ile	AUC	AUC	AUC	AUC
Leu	CUU, CUC, CUG	CUU, CUC, CUG	CUG	CUG
Val	GUC	GUC, GUG	GUG	GUG
Arg	CGU, CGC, AGG	CGU, CGC	CGC	CGC

Amino Acid	*L. seymouri*	*T. cruzi*	*T. theileri*	*T. brucei*
Thr	ACU, ACA	ACU, ACC, ACA	ACC, ACG	ACG
Ala	GCA	GCU, GCC, GCA	GCC, GCG	GCC
Pro	CCU, CCA	CCU, CCA	CCC, CCG	CCC
Ser	UCU, UCA, AGU	UCU, UCC, UCA, AGU	UCC, UCG, AGC	UCC
Ile	AUU, AUA	AUU	AUC	AUC
Leu	UUG	CUU, CUC	CUC, CUG	CUC
Val	GUA	GUU, GUC, GUA	GUC, GUG	GUG
Arg	CGU, CGA	CGU, CGC	CGC	CGC

Table 13 | Optimal Codons paired with the Major tRNA Genes for the Eight Excavates. Optimal codons that complement the major tRNA gene are written in blue. The optimal codons that complement the major tRNA gene if the adenosine at the wobble position underwent deamination to inosine are written in green.

4.6.4 Alveolata

The only tRNA gene which complements any of the ADAT preferred codons found within the alveolates was that of $\text{tRNA}^{\text{Ile}}_{\text{GAT}}$ in *S. coeruleus*. This was unusual due to the presence of nine genes for $\text{tRNA}^{\text{Ile}}_{\text{AAT}}$, the co-existence of both the G34 and A34 gene is not usually seen within Eukaryota (Grosjean, 2009). This could be due to an error by tRNA-scan, however, there is no indication that this is the case.

As previously mentioned no TAPSILVR tRNA genes within any of the *Plasmodium* species could be deemed major as the number of genes for the tRNA molecules were identical. The majority of the A34 tRNA genes were found to be major within *T. thermophila* and *S. coeruleus*, with the exception of $tRNA^{Pro}_{AGG}$ in *S. coeruleus* and $tRNA^{Arg}_{ACG}$ in both species. Therefore, due to the majority of the ADAT preferred codons being optimal, the data was consistent with the presence of the I34 tRNA deamination. Furthermore, when the deamination is taken into account, the percentage that the optimal codons match the major tRNA genes increases from 41% to 78% (Table 14). Hence, within the alveolates the I34 tRNA deamination appears to allow for the selection of the ADAT preferred codons. The reasons behind this selectional force can be clearly seen in *T. thermophila*, *S. coeruleus* and *P. falciparum*, where the ADAT preferred codons are seen at a significantly higher proportion within their HEG when compared to their WEG, showing selection via translational efficiency (Table 7). The ADAT preferred codons are only shown to be selected due to translational accuracy within *T. thermophila* where they are collectively found in significantly higher proportions within the protein coding domains than the non-domains (Table 8).

Amino Acid	*T. thermophila*	*S. coeruleus*
Thr	ACU, ACC	ACU, ACC
Ala	GCU, GCC	GCU, GCC
Pro	CCU, CCC	CCA
Ser	UCU, UCC	UCA
Ile	AUU, AUC	AUU, AUC
Leu	UUG, CUU, CUC	UUG, CUU, CUC, CUG
Val	GUU, GUC	GUC
Arg	CGU, AGA	CGU, AGA

Table 14 | Optimal Codons paired with the Major tRNA Genes for the Alveolates *T. thermophila* and *S. coeruleus*. Optimal codons that complement the major tRNA gene are written in blue. The optimal codons that complement the major tRNA gene if the adenosine at the wobble position underwent deamination to inosine are written in green.

4.6.5 Stramenopiles

The majority of the G34 tRNA genes were absent from the stramenopiles with the exceptions of: $tRNA^{Ala}_{GGC}$, $tRNA^{Arg}_{GCG}$ and $tRNA^{Ile}_{GAT}$ within *P. infestans*, $tRNA^{Ile}_{GAT}$ within *P. halstedii* and, most notably, $tRNA^{Val}_{GAC}$ within *A. astaci*, *A. invadans*, *S. parasitica* and *N. gaditana*. Despite this, the ADAT preferred codon for alanine was deemed optimal in all of the stramenopiles, with the other ADAT preferred codons being found to be optimal in up to seven of the stramenopiles.

The ADAT preferred codon for valine was only found to be optimal in *A. invadans* and *S. parasitica*, were its corresponding tRNA gene was found to be optimal. Within the stramenopiles, the A34 tRNA molecule for valine does not appear to deaminate and if it does deaminate it appears to have no effect on the selection for the ADAT preferred codon of valine.

The A34 tRNA genes within the stramenopiles were often found to be the major tRNA genes for their respective amino acids. This, when the I34 tRNA deamination was considered, affected whether or not the optimal codons of the TAPSILVR amino acids complemented the major tRNA genes. Without the presence of the deamination the optimal codons found for the TAPSILVR amino acids complemented 19% of the major tRNA genes, however, with the deamination this percentage rises to 39% (Table 15). Therefore, the I34 tRNA deamination allows for the selection for the ADAT preferred codons, with the exception of that for valine, within the stramenopiles.

The ADAT preferred codons are seen in significantly higher frequencies within the HEG than the WEG in the stramenopiles with the exceptions of *P. multistriata* and *N. gaditana* (Table 7). Therefore, the selection for the ADAT preferred codons within the majority of the stramenopiles is due to translational efficiency. The ADAT preferred codons appear to be selected for within the genomes of *S. parasitica*, *N. gaditana* and *A. anophagefferens* due to translational accuracy. This can be seen as collectively the ADAT preferred codons were found at significantly higher frequencies within their protein coding domains than their non-domains (Table 8).

Amino Acid	***P. infestans***	***P. sojae***	***P. halstedii***	***A. asaci***	***A. invadans***
Thr	ACC, ACG	ACC, ACG	ACG	ACC, ACG	ACG
Ala	GCC	GCC	GCU, GCC	GCC, GCG	GCC, GCG
Pro	CCC, CCG	CCC, CCG	CCU	CCC, CCG	CCC, CCG
Ser	UCC, UCG, AGC	UCC, UCG, AGC	UCG	UCG	UCG
Ile	AUC	AUC	AUU	AUC	AUC
Leu	CUC, CUG	CUC, CUG	CUU, CUG	UUG, CUC	UUG, CUC
Val	GUG	GUG	GUG	GUG	GUC, GUG
Arg	CGU, CGC	CGC	CGU, CGC	CGC	CGC
Amino Acid	***S. parasitica***	***P. multistriata***	***T. pseudonana***	***N. gaditana***	***A. anophagefferans***
Thr	ACG	ACU	ACC	ACU	ACU, ACC, ACA
Ala	GCU, GCC	GCU, GCA	GCC	GCU	GCU, GCC, GCA
Pro	CCC	CCU, CCA	CCC	CCU	CCU, CCC, CCA
Ser	UCG	UCU, UCA, AGU	UCC	UCU, UCA, AGU	UCU, UCA, AGU
Ile	AUC	AUU	AUC	AUU	AUU, AUA
Leu	UUA, CUC	UUA, UUG, CUU	UUA, CUC	UUA, UUG, CUU	CUU, CUA
Val	GUC	GUU, GUA	-	GUU, GUA	GUU, GUA
Arg	CGU, CGC	CGU, CGA, AGA	CGU	CGU, CGC, CGA, AGA	CGU, CGC, AGA, AGG

Table 15 | Optimal Codons paired with the Major tRNA Genes for the Ten Stramenopiles. Optimal codons that complement the major tRNA gene are written in blue. The optimal codons that complement the major tRNA gene if the adenosine at the wobble position underwent deamination to inosine are written in green.

4.6.6 Rhizaria

The ADAT preferred codons for the SILVR amino acids were found to be optimal in both rhizarians, whilst the ADAT preferred codon for threonine was found optimal in *B. natans* and the ADAT preferred codon for alanine was found optimal in *P. brassicae*. No tRNA genes were found for those tRNA molecules that complement the ADAT preferred codons. However, consistent with the I34 tRNA deamination, the A34 tRNA genes were present and, in most cases, major.

When compared to the major tRNA genes the optimal codons for both rhizarians matched 38% of the time, however, when the I34 deamination was considered this increased to 58% (Table 16). Therefore, the deamination of A34 to I34 in Rhizaria is shown to contribute to the selection for the

ADAT preferred codons. The selection for the ADAT preferred codons appears to be due to both translational efficiency and accuracy. As both species showed significantly higher levels of ADAT preferred codons within their HEG when compared to their WEG (Table 7), and again in their protein coding domains when compared to their non-domains (Table 8).

Amino Acid	*B. natans*	*P. brassicae*
Thr	ACU, ACC	ACG
Ala	GCU	GCC, GCG
Pro	CCU	CCG
Ser	UCU, UCC	UCG
Ile	AUC	AUC
Leu	CUU, CUC, CUA	CUG
Val	GUU, GUA	GUC
Arg	CGU, CGC	CGU, CGC

Table 16 | Optimal Codons paired with the Major tRNA Genes for the Two Rhizarians. Optimal codons that complement the major tRNA gene are written in blue. The optimal codons that complement the major tRNA gene if the adenosine at the wobble position underwent deamination to inosine are written in green.

4.7 Presence of the I34 tRNA deamination

The modified versions of all the TAPSILVR tRNA screened in *S. arctica* (opisthokont), *G. lambila* (excavate), *P. falciparum* (alveolate), *P. sojae* (stramenopile) and *T. pseudonana* (stramenopile) were present within the nc/miRNA reads (Table 9). In *D. discoideum*, *C. owczarzaki*, *L. major* and *P. infestans* the modified versions of the TAPSILVR tRNA were found for some amino acids but for others no modified version could be found (Table 9). Therefore, evidence that the I34 tRNA deamination occurs was found for all of the eukaryotic supergroups examined, except that for Rhizaria.

The amoebozoan *E. histolytica* and the excavate *T. brucei* showed no evidence for the presence of the A34-to-I34 modified tRNA molecules. However, as unmodified tRNA genes were also not located within the nc/miRNA reads, despite the presence of their genes within the organisms' genome, it suggests that the screen was too small due to a lack of nc/small RNA data.

4.8 Presence of the A37-to-I37 deamination

The presence of the I37 modification without the I34 modification was not seen in any of the protists. The I37 modification was present when the I34 modification was also present in *D. discoideum* (amoebozoan), *P. sojae* (stramenopile) and *T. pseudonana* (stramenopile) (Table 10). Therefore, this study provides evidence that the I37 modification carried out by ADAT1 is dependent on the I34 deamination carried out by hetADAT. Further research is required to determine the validity of this observation.

4.9 The Selection towards ADAT preferred codons throughout the Supergroups

The 33 protists researched within this study varied greatly when it came to their CUB. Despite this, the majority of the protists showed evidence that the I34 tRNA deamination allowed for the selection of the ADAT preferred codons. When the frequency of the ADAT preferred codons for the TAPSILVR amino acids were viewed collectively within the HEG and the WEG, 28 of the organisms showed a significantly higher frequency of the codons within the HEG, indicating selection via translational efficiency. The frequency of the ADAT preferred codons was also viewed collectively within the protein coding domains and the non-domains of the species, here 22 of the protists showed a significantly higher frequency for the use of the codons within the protein coding domains, indicating selection via translational accuracy. Of the protists studied 23 did not have any tRNA genes present in their genome that complemented the ADAT preferred codons. Whilst the other ten protists showed the presence of only one G34 tRNA type for the TAPSILVR amino acids (with the exception of *P. infestans*). Therefore, there is a strong indication that the I34 tRNA deamination takes effect in all the organisms studied.

The tRNA gene tRNA$^{Val}_{GAC}$ appears in seven of the protists studied, *E. histolytica*, *G. lamblia*, *S. salmonicida*, *P. sojae*, *A. astaci*, *A. invadans* and *N. gaditana*, whilst in these same organisms the presence of the tRNA$^{Val}_{AAC}$ gene could not be found. There is a possibility that hetADAT enzyme lost the ability to deaminate the A34 tRNA molecule for valine within these organisms, however, due to this pattern appearing throughout the supergroups, this loss would have had to happen via convergent evolution.

Within the Amoebozoa, Opisthokonta and Apusomonada the ADAT preferred optimal codons were almost always found to compliment the major tRNA gene, A34 (86%). This percentage decreases to 63% for the excavates, 61% for the alveolates, 56% for the stramenopiles and only 33% for the rhizarians. Therefore, it appears that the deamination of A34 TAPSILVR tRNA within the Amoebozoa, Opisthokonta and the Apusomonada allows for a higher rate of selection of the ADAT preferred codons, than those supergroups found on the other side of the tree (Figure 3).

Due to the findings of this study it can be concluded that the occurrence of the I34 tRNA deamination allows for the selection of the ADAT preferred codons within the seven eukaryotic supergroups observed in this study. Therefore, this suggests that this occurrence is either a result of two to seven cases of convergent evolution or that an unknown common ancestor of the supergroups held the ability to deaminate A34-to-I34 on TAPSILVR tRNA genes leading towards the selective advantage of the ADAT preferred codons.

As the I34 deamination has only been found to affect arginine tRNA molecules within Bacteria and has not been observed in Archaea (Gerber, 1999), the mechanism behind the TAPSILVR I34 tRNA deamination must have evolved within the Eukaryota. This study revealed that the deamination evolved within a eukaryotic protist ancestral to the known eukaryotic supergroups. It is likely that within this eukaryotic ancestor the gene for ADAT1 duplicated and diverged into the genes for ADAT2 and ADAT3, the subunits of hetADAT (Gerber, 1999). Therefore, further research into the evolution of the ADAT1 enzymes and their effect on tRNA molecules (the A37-to-I37 modification) must be carried out in order to fully understand the mechanisms behind the A34 tRNA deamination within the eukaryotes.

4.10 Conclusion

The analysis of the CUB in the early eukaryotic protists studied revealed the presence of a selective pressure toward ADAT preferred codons within seven of the eukaryotic supergroups. Due to the wide spread absence of the complimentary tRNA genes for the ADAT preferred codons whilst the most abundant tRNA gene for each TAPSILVR amino acid is often found to be those with adenosine at their wobble position. This study, therefore provides evidence that the origin of A34-to-I34 modification carried out by the hetADAT enzyme lies in the last common ancestor of the eukaryotes, where it most likely led to the selection of the ADAT preferred codons.

5.0 References

Akashi, H. (1994). Synonymous codon usage in *Drosophila melanogaster*: natural selection and translational accuracy. Genetics, 136(3):927-935.

Akashi, H., & Eyre-Walker, A. (1998). Translational selection and molecular evolution. Current Opinion in Genetics & Development, 8(6), 688–693. doi:10.1016/s0959-437x(98)80038-5

Altschul, S.F., Gish, W., Miller, W., Meyers, E.W., Lipman, D.J. (1990). Basic local alignment search tool. Journal of Molecular Biology, 215(3), 403-410. doi:10.1006/jmbi.1990.9999; 10.1016/S0022-2836(05)80360-2

Bapteste, E., Brinkmann, H., Lee, J. A., Moore, D. V., Sensen, C. W., Gordon, P., … Philippe, H. (2002). The analysis of 100 genes supports the grouping of three highly divergent amoebae: Dictyostelium, Entamoeba, and Mastigamoeba. Proceedings of the National Academy of Sciences, 99(3), 1414–1419. doi:10.1073/pnas.032662799

Berriman, M. (2005). The Genome of the African Trypanosome *Trypanosoma brucei*. Science, 309(5733), 416–422. doi:10.1126/science.1112642

Boccaletto, P., Machnicka, M. A., Purta, E., Piątkowski, P., Bagiński, B., Wirecki, T. K., Crecy-Lagard, V., Ross, R., Limbach, P. A., Kotter, A., Helm, M., Bujnicki, J. M. (2017). MODOMICS: a database of RNA modification pathways. 2017 update. Nucleic Acids Research, 46(D1), D303–D307. doi:10.1093/nar/gkx1030

Bramall, E. (2019). Investigating Retrotransposons in the Close Relatives of Animals. Undergraduate Dissertation. The University of Huddersfield.

Brown, M. W., Spiegel, F. W., & Silberman, J. D. (2009). Phylogeny of the “Forgotten” Cellular Slime Mold, *Fonticula alba*, Reveals a Key Evolutionary Branch within Opisthokonta. Molecular Biology and Evolution, 26(12), 2699–2709. doi:10.1093/molbev/msp185

Brown, M. W., Heiss, A. A., Kamikawa, R., Inagaki, Y., Yabuki, A., Tice, A. K., Shiratori, T., Ishida, K., Hashimoto, T., Simpson, A. G., Roger, A. J. (2018). Phylogenomics Places Orphan Protistan Lineages in a Novel Eukaryotic Super-Group. Genome Biology and Evolution, 10(2), 427-233. doi:10.10.93/gbe/evy014.

Burki, F., & Keeling, P. J. (2014). Rhizaria. Current Biology, 24(3), R103–R107. doi:10.1016/j.cub.2013.12.025

Burki, F., Roger, A. J., Brown, M. W., & Simpson, A. G. B. (2019). The New Tree of Eukaryotes. Trends in Ecology & Evolution. doi:10.1016/j.tree.2019.08.008

Cavalier-Smith, T., & Chao, E. E. (2010). Phylogeny and Evolution of Apusomonadida (Protozoa: Apusozoa): New Genera and Species. Protist, 161(4), 549–576. doi:10.1016/j.protis.2010.04.002

Chan, P. P., & Lowe, T. M. (2009). GtRNAdb: a database of transfer RNA genes detected in genomic sequence. Nucleic Acids Research, 37(Database), D93–D97. doi:10.1093/nar/gkn787

Comeron, J. M., & Aguadé, M. (1998). An Evaluation of Measures of Synonymous Codon Usage Bias. Journal of Molecular Evolution. 47(3), 268–274. doi:10.1007/pl00006384

Crick, F.H.C. (1966). Codon-anticodon pairing: The wobble hypothesis. Journal of Molecular Biology. 19(2), 548–555. doi:10.1016/s0022-2836(66)80022-0

Del Campo, J., Sieracki, M. E., Molestina, R., Keeling, P., Massana, R., & Ruiz-Trillo, I. (2014). The others: our biased perspective of eukaryotic genomes. Trends in Ecology & Evolution, 29(5), 252–259. doi:10.1016/j.tree.2014.03.006

Deng, N., Liu, C., Song, Q., Yao, M., Yang, R., Ma, F., & Tian, Y. (2020). Codon usage pattern analysis of *Gnetum luofuense* (Gnetales) based on transcriptome data. Agronomy Journal. doi:10.1002/agj2.20383

Duret, L., & Mouchiroud, D. (1999). Expression pattern and, surprisingly, gene length shape codon usage in Caenorhabditis, Drosophila, and Arabidopsis. Proceedings of the National Academy of Sciences, 96(8), 4482–4487. doi:10.1073/pnas.96.8.4482

Eichinger, L., Pachebat, J. A., Glockner, G., Rajandream, M. A., Sucgang, R., Berriman, M., Song, J., Olsen, R., Szafranski, K., Xu, Q., Tunggal, B., Kummerfeld, S., Madera, M., Konfortov, B. A., Rivero, F., Bankier, A. T., Lehmann, R., Hamlin, N., Davies, R., Gaudet, P., Fay, P., Pilcher, K., Chen, G., Saunders, D., Sodergren, E., Davis, P., Kerhornou, A., Nie., X., Hall, N., Anjard, C., Hemphill, L., Bason, N., Farbrother, P., Desany, B., Just, E., Morio, T., Rost, R., Churcher, C., Cooper, J., Haydock, S., Van Driessche, N., Cronin, A., Goodhead, I., Muzny, D., Mourier, T., Pain, A., Lu, M., Harper, D., Lindsay, R., Hauser, H., James, K., Quiles, M., Madan Babu, M., Saito, T., Buchrieser, C., Wardroper, A., Felder, M., Thangavelu, M., Johnson, D., Knights, A., Loulseged, H., Mungall, K., Oliver, K., Price, C., Quail, M. A., Urushihara, H., Hernandez, J., Rabbinowitsch, E., Steffen, D., Sanders, M., Ma, J., Kohara, Y., Sharp, S., Simmonds, M., Spiegler, S., Tivey, A., Sugano, S., White, B., Walker, D., Woodward, J., Winckler, T., Tanaka, Y., Shaulsky, G., Schleicher, M., Weinstock, G., Rosenthal, A., Cox, E. C., Chisholm, R. L., Gibbs, R., Loomis, F. W., Platzer, M., Kay, R. R., Williams, J., Dear, P. H., Nogel, A. A., Barrell, B, & Kuspa, A. (2005). The genome of the social amoeba *Dictyostelium discoideum*. Nature, 435(7038), 43–57. doi:10.1038/nature03481

Einarsson, E., Ma'ayeh, S., & Svärd, S. G. (2016). An up-date on Giardia and giardiasis. Current Opinion in Microbiology, 34, 47–52. doi:10.1016/j.mib.2016.07.019

El-Sayed, N. M. (2005). The Genome Sequence of Trypanosoma cruzi, Etiologic Agent of Chagas Disease. Science, 309(5733), 409–415. doi:10.1126/science.1112631

Figuet, E., Ballenghien, M., Romiguier, J., & Galtier, N. (2014). Biased Gene Conversion and GC-Content Evolution in the Coding Sequences of Reptiles and Vertebrates. Genome Biology and Evolution, 7(1), 240–250. doi:10.1093/gbe/evu277

Flegontov, P., Butenko, A., Firsov, S., Kraeva, N., Eliáš, M., Field, M. C., Filatov, D., Flegontova, O., Gerasimov, E. S., Hlaváčová, J., Ishemgulova, A., Jackson, A. P., Kelly, S., Kostygov, A. Y., Logacheva, M. D., Maslov, D. A., Opperdoes, F. R., O'Reilly, A., Sádlová, J., Ševčíková, T., Venkatesh, D., Vlček, Č., Volf, P., Votýpka, J., Záhonová, K., Yurchenko, V., Lukeš, J. (2016). Genome of *Leptomonas pyrrhocoris*: a high-quality reference for monoxenous trypanosomatids and new insights into evolution of *Leishmania*. Scientific Reports. Article number: 23704 (2016).

Gasteiger E., Gattiker A., Hoogland C., Ivanyi I., Appel R.D., Bairoch A. ExPASy: the proteomics server for in-depth protein knowledge and analysis. Nucleic Acids Res. 31:3784- 3788 (2003).

Gerber, A. P. (1999). An Adenosine Deaminase that Generates Inosine at the Wobble Position of tRNAs. Science, 286(5442), 1146–1149. doi:10.1126/science.286.5442.1146

Gouy, M., & Gautier, C. (1982). Codon usage in bacteria: correlation with gene expressivity. Nucleic Acids Research, 10(22), 7055–7074. doi:10.1093/nar/10.22.7055

Grantham, R., Gautier, C., Gouy, M., Mercier, R., & Pavé, A. (1980). Codon catalog usage and the genome hypothesis. Nucleic Acids Research. 8(1), 197–197. 10.1093/nar/8.1.197-c

Grosjean, H., de Crécy-Lagard, V., & Marck, C. (2009). Deciphering synonymous codons in the three domains of life: Co-evolution with specific tRNA modification enzymes. FEBS Letters, 584(2), 252–264. doi:10.1016/j.febslet.2009.11.052

Gupta, S., & Ghosh, T. . (2001). Gene expressivity is the main factor in dictating the codon usage variation among the genes in *Pseudomonas aeruginosa*. Gene, 273(1), 63–70. doi:10.1016/s0378-1119(01)00576-5

Haas, B. J., Kamoun, S., Zody, M. C., Jiang, R. H. Y., Handsaker, R. E., Cano, L. M., … Torto-Alalibo, T. (2009). Genome sequence and analysis of the Irish potato famine pathogen *Phytophthora infestans*. Nature, 461(7262), 393–398. doi:10.1038/nature08358

Jukes, T. H., & Osawa, S. (1993). Evolutionary changes in the genetic code. Comparative Biochemistry and Physiology Part B: Comparative Biochemistry, 106(3), 489–494. doi:10.1016/0305-0491(93)90122-l

Iben, J. R., & Maraia, R. J. (2012). tRNAomics: tRNA gene copy number variation and codon use provide bioinformatic evidence of a new anticodon:codon wobble pair in a eukaryote. RNA, 18(7), 1358–1372. doi:10.1261/rna.032151.111

Ikemura, T. (1981). Correlation between the abundance of *Escherichia coli* transfer RNAs and the occurrence of the respective codons in its protein genes: A proposal for a synonymous codon choice that is optimal for the *E. coli* translational system. Journal of Molecular Biology, 151(3), 389–409. doi:10.1016/0022-2836(81)90003-6

Ivens, A. C. (2005). The Genome of the Kinetoplastid Parasite, *Leishmania major*. Science, 309(5733), 436–442. doi:10.1126/science.1112680

Keeling, P. J., Burger, G., Durnford, D. G., Lang, B. F., Lee, R. W., Pearlman, R. E., … Gray, M. W. (2005). The tree of eukaryotes. Trends in Ecology & Evolution, 20(12), 670–676. doi:10.1016/j.tree.2005.09.005

Kersey, P.J., Allen, J.E., Allot, A., Barba, M., Boddu, S., Bolt, B.J., Carvalho-Silva, D., Christensen, M., Davis, P., Grabmueller, C., Kumar, N., Liu, Z., Maurel, T., Moore, B., McDowall, M.D., Maheswari, U., Naamati, G., Newman, V., Ong, C.K., Paulini, M., Pedro, H., Perry, E., Russell, M., Sparrow, H., Tapanari, E., Taylor, K., Vullo, A., Williams, G., Zadissia, A., Olson, A., Stein, J., Wei, S., Tello-Ruiz, M., Ware, D., Luciani, A., Potter, S., Finn, R.D., Urban, M., Hammond-Kosack, K.E., Bolser, D.M., De Silva, N., Howe, K.L., Langridge, N., Maslen, G., Staines, D.M. & Yates, A. (2018). Ensembl genomes 2018: An integrated omics infrastructure for non-vertebrate species. Nucleic Acids Research, 46(1), D808. doi:10.1093/nar/gkx1011tRNAscan-SE

Kraeva, N., Butenko, A., Hlaváčová, J., Kostygov, A., Myškova, J., Grybchuk, D., Leštinová, T., Votýpka, J., Volf, P., Opperdoes, F., Flegontov, P., Lukeš, J., Yurchenko, V. (2015). Leptomonas seymouri: Adaptations to the Dixenous Life Cycle Analyzed by Genome Sequencing, Transcriptome Profiling and Co-infection with *Leishmania donovani*. PLOS Pathogens, 11(8), e1005127. doi:10.1371/journal.ppat.1005127

Lahr, D. J. G., Bergmann, P. J., & Lopes, S. G. B. C. (2008). Taxonomic Identity in Microbial Eukaryotes: A Practical Approach Using the Testate Amoeba *Centropyxis* to Resolve Conflicts

Between Old and New Taxonomic Descriptions. Journal of Eukaryotic Microbiology, 55(5), 409–416. doi:10.1111/j.1550-7408.2008.00339.x

Lara, E., Heger, T. J., Ekelund, F., Lamentowicz, M., & Mitchell, E. A. D. (2008). Ribosomal RNA Genes Challenge the Monophyly of the Hyalospheniidae (Amoebozoa: Arcellinida). Protist, 159(2), 165–176. doi:10.1016/j.protis.2007.09.003

Leander, B. S., Kuvardina, O. N., Aleshin, V. V., Mylnikov, A. P., & Keeling, P. J. (2003). Molecular Phylogeny and Surface Morphology of *Colpodella edax* (Alveolata): Insights into the Phagotrophic Ancestry of Apicomplexans. The Journal of Eukaryotic Microbiology, 50(5), 334–340. doi:10.1111/j.1550-7408.2003.tb00145.x

Li, H. (2013). Seqtk: a fast and lightweight tool for processing FASTA or FASTQ sequences. https://github.com/lh3/seqtk

Lim, V. I. (1995). Analysis of Action of the Wobble Adenine on Codon Reading within the Ribosome. Journal of Molecular Biology, 252(3), 277–282. doi:10.1006/jmbi.1995.0494

Lim, V. I., Curran, J. F. (2001). Analysis of codon:anticodon interactions within the ribosome provides new insights into codon reading and the genetic code structure. RNA. Jul;7(7):942-57. doi: 10.1017/s135583820100214x.

Liu, X.-Y., Li, Y., Ji, K.-K., Zhu, J., Ling, P., Zhou, T., Fan, L-Y., Xie, S.-Q. (2020). Genome-wide codon usage pattern analysis reveals the correlation between codon usage bias and gene expression in *Cuscuta australis*. Genomics. doi:10.1016/j.ygeno.2020.03.002

Loftus, B., Anderson, I., Davis, R., Alsmark, U. C. M., Samuelson, J., Amedeo, P., Roncaglia, P., Berriman, M., Hirt, R. P., Mann, B. J., Nozaki, T., Suh, B., Pop, M., Duchene, M., Ackers, J., Tannich, E., Leippe, M., Hofer, M., Bruchhaus, I., Willhoeft, U., Bhattacharya, A., Chillingworth, T., Churcher, C., Hance, Z., Harris, B., Harris, D., Jagels, K., Moule, S., Mungall, K., Ormond, D., Squares, R., Whitehead, S., Quail, M. A., Rabbinowitsch, E., Norbertczak, H., Price, C., Wang, Z., Guillen, N., Gilchrist, C., Stroup, S. E., Bhattacharya, S., Lohia, A., Foster, P. G., Sicheritz-Ponten, T., Weber, C., Singh, U., Mukherjee, C., El-Sayed, N, M., Petri Jr, W. A., Clark, G., Embley, M., Barrell, B., Fraser, C. M., Hall, N. (2005). The genome of the protist parasite *Entamoeba histolytica*. Nature, 433(7028), 865–868. doi:10.1038/nature03291

Lowe, T.M., Chan, P.P. (2016). tRNAscan-SE On-line: Search and Contextual Analysis of Transfer RNA Genes. Nucl. Acids Res. 44: W54-57

Lowe,T.M. and Eddy,S.R. (1997) tRNAscan-SE: a program for improved detection of transfer RNA genes in genomic sequence. Nucleic Acids Res., 25, 955–964.

Maas, S., Gerber, A. P., & Rich, A. (1999). Identification and characterization of a human tRNA-specific adenosine deaminase related to the ADAR family of pre-mRNA editing enzymes. Proceedings of the National Academy of Sciences, 96(16), 8895–8900. doi:10.1073/pnas.96.16.8895

Machado, H. E., Lawrie, D. S., & Petrov, D. A. (2019). Pervasive Strong Selection at the Level of Codon Usage Bias in *Drosophila melanogaster*. Genetics, genetics.302542.2019. doi:10.1534/genetics.119.302542

Marchler-Bauer, A., Bo, Y., Han, L., He, J., Lanczycki, C.J., Lu, S., Chitsaz, F., Derbyshire, M.K., Geer, R.C., Gonzales, N.R., Gwadz, M., Hurwitz, D.I., Lu, F., Marchler, G.H., Song, J.S., Thanki, N. (2017). CDD/SPARCLE: Functional classification of proteins via subfamily domain architectures. Nucleic Acids Research, 45(1), D203. doi:10.1093/nar/gkw1129

Marck, C., & Grosjean, H. (2002). tRNomics: Analysis of tRNA genes from 50 genomes of Eukarya, Archaea, and Bacteria reveals anticodon-sparing strategies and domain-specific features. RNA, 8(10), 1189–1232. doi:10.1017/s1355838202022021

Metzl-Raz, E., Kafri, M., Yaakov, G., Barkai, N. 2020. Gene transcription as a limiting factor in protein production and cell growth. bioRxiv. doi: https://doi.org/10.1101/626242

Milne, I., Stephen, G., Bayer, M., Cock, P.J.A., Pritchard, L., Cardle, L., Shaw, P.D., Marshall, D. (2013). Using Tablet for visual exploration of second-generation sequencing data. Briefings in Bioinformatics 14(2), 193-202.

Moriyama, E. (1998). Gene length and codon usage bias in *Drosophila melanogaster*, Saccharomyces cerevisiae and Escherichia coli. Nucleic Acids Research, 26(13), 3188–3193. doi:10.1093/nar/26.13.3188

Musto, H., Naya, H., Zavala, A., Romero, H., Alvarez-Valín, F., & Bernardi, G. (2006). Genomic GC level, optimal growth temperature, and genome size in prokaryotes. Biochemical and Biophysical Research Communications, 347(1), 1–3. doi:10.1016/j.bbrc.2006.06.054

Novoa, E. M., Pavon-Eternod, M., Pan, T., & Ribas de Pouplana, L. (2012). A Role for tRNA Modifications in Genome Structure and Codon Usage. Cell, 149(1), 202–213. doi:10.1016/j.cell.2012.01.050

Ortiz, S., Solari, A. (2019). Excavata-Kinetoplastea Trypanosomatidae Parasites and the Interaction with their Hosts. IJoTD. doi:10.23937/IJTD-2017/1710015

Pakula, T. M. (2005). The effect of specific growth rate on protein synthesis and secretion in the filamentous fungus *Trichoderma reesei*. Microbiology, 151(1), 135–143. doi:10.1099/mic.0.27458-0

Pawlowski, J., & Burki, F. (2009). Untangling the Phylogeny of Amoeboid Protists. Journal of Eukaryotic Microbiology, 56(1), 16–25. doi:10.1111/j.1550-7408.2008.00379.x

Peden, J. F. (1999). Analysis of Codon Usage. Department of Genetics. Thesis submitted to the University of Nottingham for the Degree of Doctor of Philosophy.

Percudani, R. (2001). Restricted wobble rules for eukaryotic genomes. Trends in Genetics, 17(3), 133–135. doi:10.1016/s0168-9525(00)02208-3

Perry,J., Dai,X. and Zhao,Y. (2005) A mutation in the anticodon of a single tRNAala is sufficient to confer auxin resistance in Arabidopsis. Plant Physiol., 139, 1284–1290.

Porath, H. T., Knisbacher, B. A., Eisenberg, E., & Levanon, E. Y. (2017). Massive A-to-I RNA editing is common across the Metazoa and correlates with dsRNA abundance. Genome Biology, 18(1). doi:10.1186/s13059-017-1315-y

Rafels-Ybern, À., Torres, A. G., Grau-Bove, X., Ruiz-Trillo, I., & Ribas de Pouplana, L. (2017). Codon adaptation to tRNAs with Inosine modification at position 34 is widespread among Eukaryotes and present in two Bacterial phyla. RNA Biology, 15(4-5), 500–507. doi:10.1080/15476286.2017.1358348

Rambaut, A. (2009). Computer program distributed by the author, website: http://tree.bio.ed.ac.uk/software/figtree/[accessed January 4, 2011]

Scala, S. (2002). Genome Properties of the Diatom *Phaeodactylum tricornutum*. Plant physiology, 129(3), 993–1002. doi:10.1104/pp.010713

Shabardina, V., Kischka, T., Kmita, H., Suzuki, Y., & Makałowski, W. (2018). Environmental adaptation of *Acanthamoeba castellanii* and *Entamoeba histolytica* at genome level as seen by comparative genomic analysis. International Journal of Biological Sciences, 14(3), 306–320. doi:10.7150/ijbs.23869

Sharp, P. M., & Devine, K. M. (1989). Codon usage and gene expression level in *Dictyostelium discoideum*: highly expressed genes do prefer optimal codons. Nucleic Acids Research, 17(13), 5029–5040. doi:10.1093/nar/17.13.5029

Sharp, P. M., Emery, L. R., & Zeng, K. (2010). Forces that influence the evolution of codon bias. Philosophical Transactions of the Royal Society B: Biological Sciences, 365(1544), 1203–1212. doi:10.1098/rstb.2009.0305

Sharp, P. M., & Li, W. (1986). An evolutionary perspective on synonymous codon usage in unicellular organisms. Journal of Molecular Evolution, 24 (1-2), 28-38. doi:10.1007/BF02099948

Sibbald, S. J., & Archibald, J. M. (2017). More protist genomes needed. Nature Ecology & Evolution, 1(5). doi:10.1038/s41559-017-0145

Simpson, A. G. ., & Roger, A. J. (2002). Eukaryotic Evolution: Getting to the Root of the Problem. Current Biology, 12(20), R691–R693. doi:10.1016/s0960-9822(02)01207-1

Sørensen, M. A., & Pedersen, S. (1991). Absolute in vivo translation rates of individual codons in *Escherichia coli*. Journal of Molecular Biology, 222(2), 265–280. doi:10.1016/0022-2836(91)90211-n

Southworth, J., Armitage, P., Fallon, B., Dawson, H., Bryk, J., & Carr, M. (2018). Patterns of Ancestral Animal Codon Usage Bias Revealed through Holozoan Protists. Molecular Biology and Evolution, 35(10), 2499–2511. doi:10.1093/molbev/msy157

Suga, H., Chen, Z., de Mendoza, A., Suga, H., Chen, Z., de Mendoza, A., Sebé-Pedrós, A., Brown, M. W., Kramer, E., Carr, M., Kerner, P., Vervoort, M., Sanchez-Pons, N., Torruella, G., Derelle, R., Manning, G., Lang, F., Russ, C., Haas, B. J., Roger, A. J., Nusbaum, C., Ruiz-Trillo, I. (2013). The Capsaspora genome reveals a complex unicellular prehistory of animals. Nature Communications, 4(1). doi:10.1038/ncomms3325

Suga, H., & Ruiz-Trillo, I. (2013). Development of ichthyosporeans sheds light on the origin of metazoan multicellularity. Developmental Biology, 377(1), 284–292. doi:10.1016/j.ydbio.2013.01.009

Torres, A. G., Piñeyro, D., Filonava, L., Stracker, T. H., Batlle, E., & Ribas de Pouplana, L. (2014). A-to-I editing on tRNAs: Biochemical, biological and evolutionary implications. FEBS Letters, 588(23), 4279–4286. doi:10.1016/j.febslet.2014.09.025

Wald, N., & Margalit, H. (2014). Auxiliary tRNAs: large-scale analysis of tRNA genes reveals patterns of tRNA repertoire dynamics. Nucleic Acids Research, 42(10), 6552–6566. doi:10.1093/nar/gku245

Wright, F. (1990). The “effective number of codons” used in a gene. Gene, 87(1), 23–29. doi:10.1016/0378-1119(90)90491-9

Xu, F., Jerlström-Hultqvist, J., Einarsson, E., Ástvaldsson, Á., Svärd, S. G., & Andersson, J. O. (2014). The Genome of *Spironucleus salmonicida* Highlights a Fish Pathogen Adapted to Fluctuating Environments. PLoS Genetics, 10(2), e1004053. doi:10.1371/journal.pgen.1004053

Zhou, Z., Dang, Y., Zhou, M., Li, L., Yu, C., Fu, J., Chen, S., Liu, Y. (2016). Codon usage is an important determinant of gene expression levels largely through its effects on transcription. Proceedings of the National Academy of Sciences, 113(41), E6117–E6125. doi:10.1073/pnas.1606724113

Zinshteyn, B., & Nishikura, K. (2009). Adenosine-to-inosine RNA editing. Wiley Interdisciplinary Reviews: Systems Biology and Medicine, 1(2), 202–209. doi:10.1002/wsbm.10

"Analyze a 2x2 contingency table”. GraphPad Software, Inc. Accessed 30 March 2020. https://www.graphpad.com/quickcalcs/contingency1.cfm

“SMALT”. Genome Research Ltd. (2015) http://www.sanger.ac.uk/resources/software/SMALT/

“2x2 Contingency Table”. VassarStats, Inc. Accessed 30 March 2020. http://vassarstats.net/tab2x2.html

6.0 Appendix

6.1 Nc Plots

General description of figures 5-37: The Nc and GC3s values of all genes found in the transcriptome for each organism are shown as grey hollow circles. The average Nc and GC3s values for each organisms set of genes are represented by a black dot. The average Nc and GC3s values for the top and bottom 5% expressed genes for each organism are represented by blue and red squares respectively. The average Nc and GC3s values for the found protein coding domains and non-domains of the top 200 expressed genes for each organisms are represented by green and purple triangles respectively. The red bell-shaped curve represents the expected position of genes if their CUB was solely influenced by a neutral mutational bias.

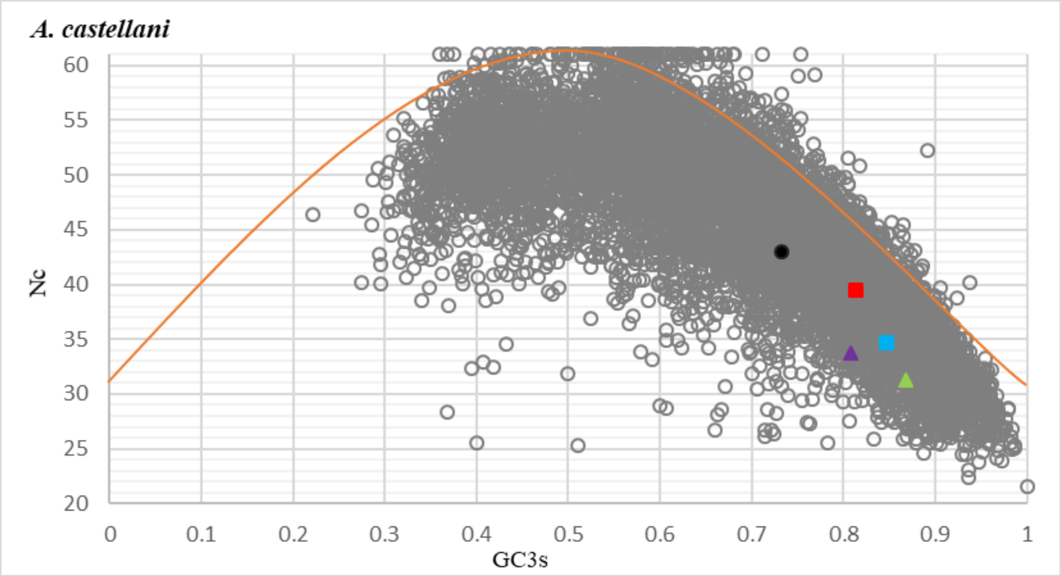


Figure 5 | Graph showing genome wide Nc plot for the Amoebozoan *A. castellani*

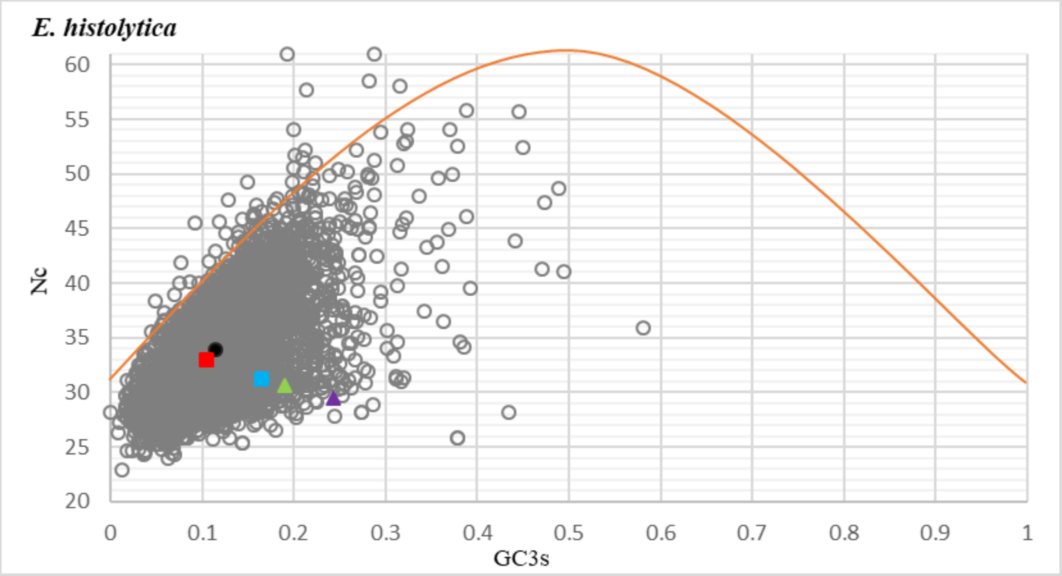


Figure 6 | Graph showing genome wide Nc plot for the Amoebozoan *E. histolytica*

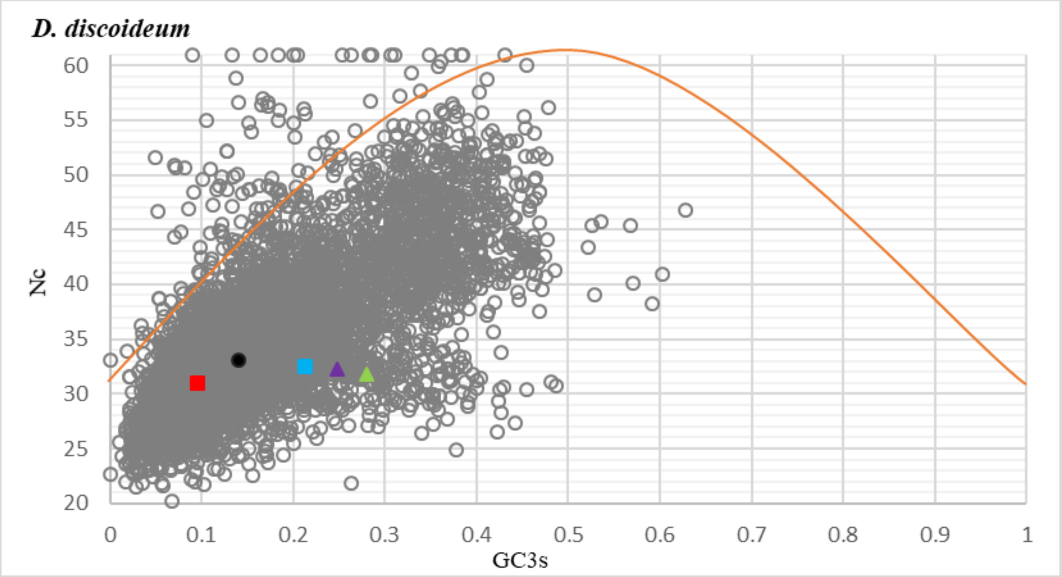


Figure 7 | Graph showing genome wide Nc plot for the Amoebozoan *D. discoideum*

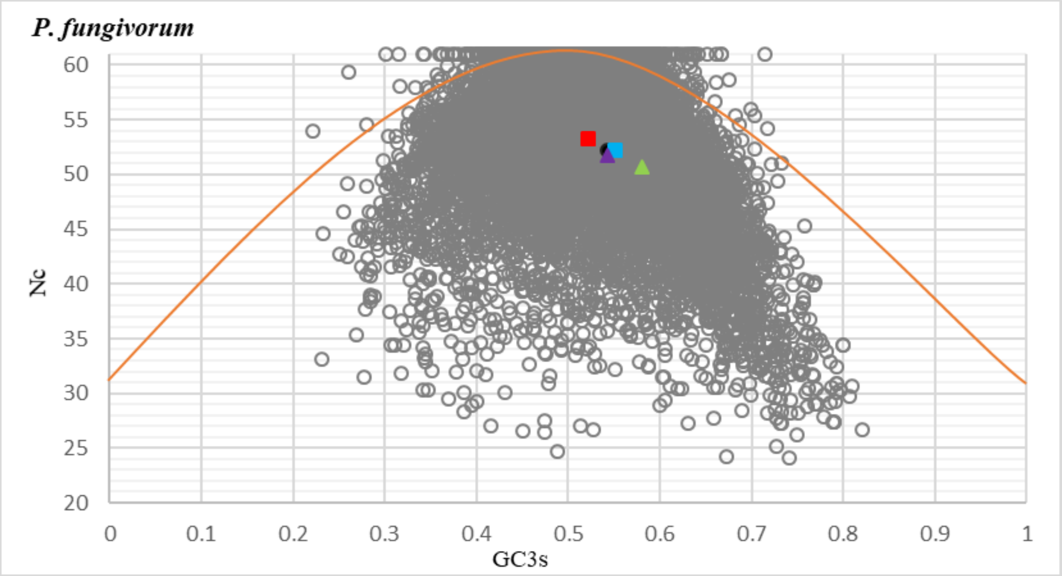


Figure 8 | Graph showing genome wide Nc plot for the Amoebozoan *P. fungivorum*

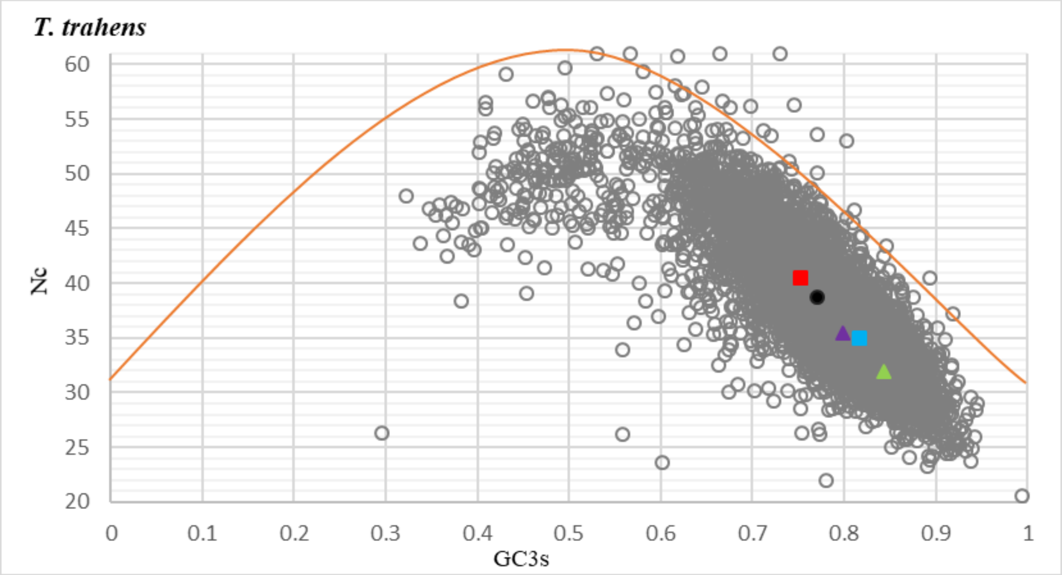


Figure 9 | Graph showing genome wide Nc plot for the Apusomonadan *T. trahens*

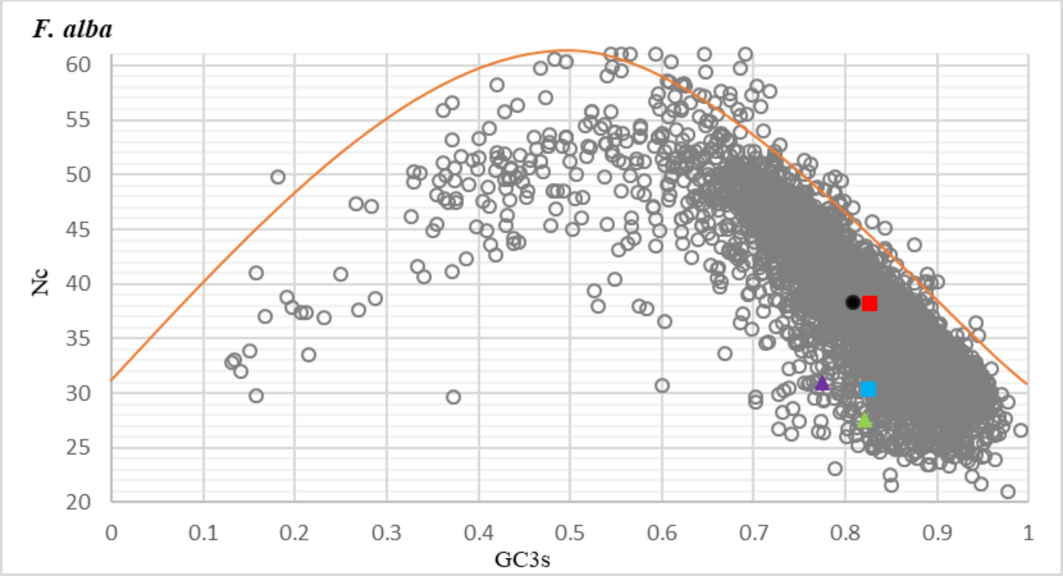


Figure 10 | Graph showing genome wide Nc plot for the Opisthokont *F. alba*

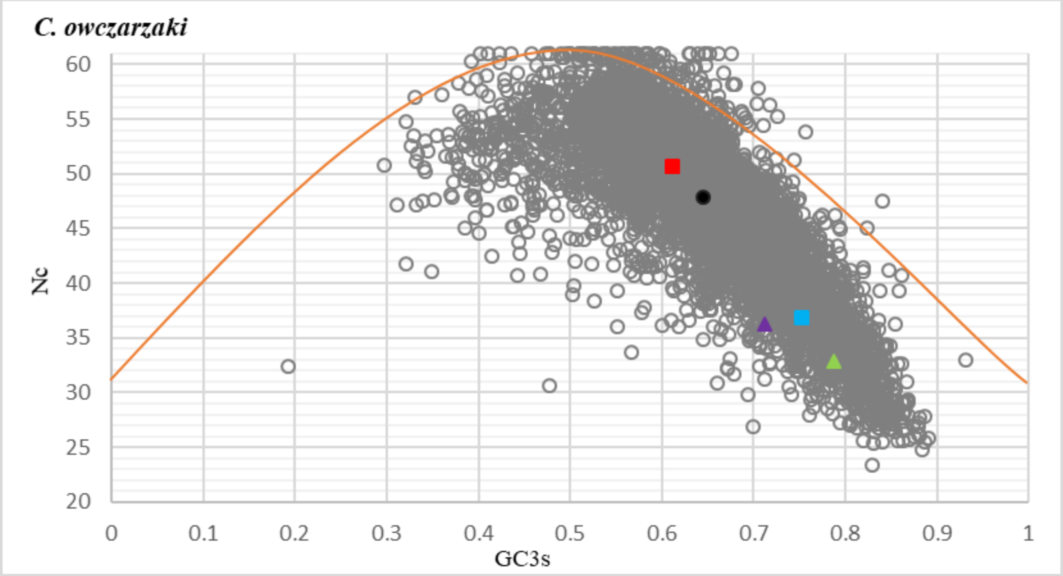


Figure 11 | Graph showing genome wide Nc plot for the Opisthokont *C. owczarzaki*

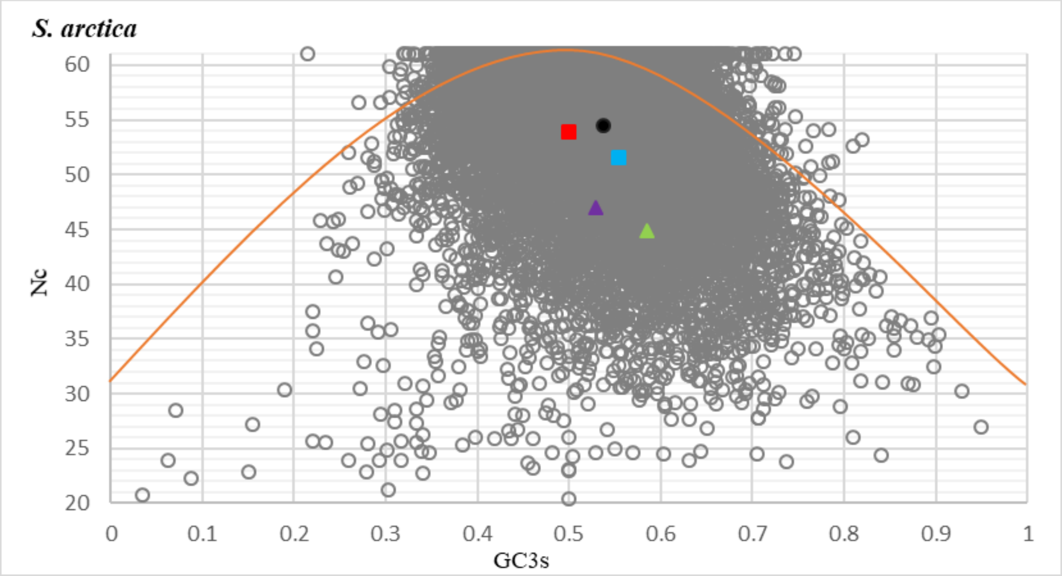


Figure 12 | Graph showing genome wide Nc plot for the Opisthokont *S. arctica*

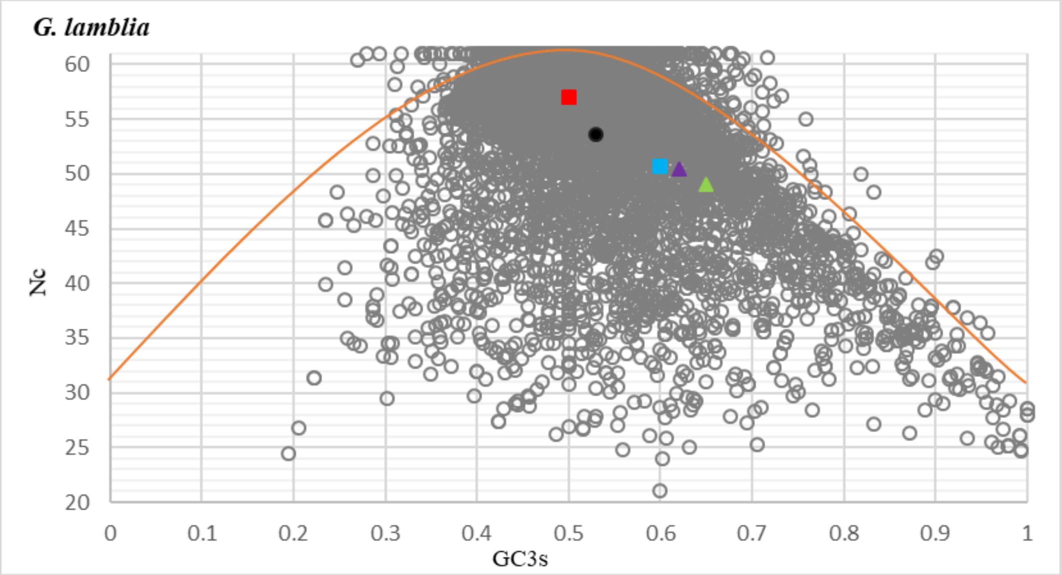


Figure 13 | Graph showing genome wide Nc plot for the Excavate *G. lamblia*

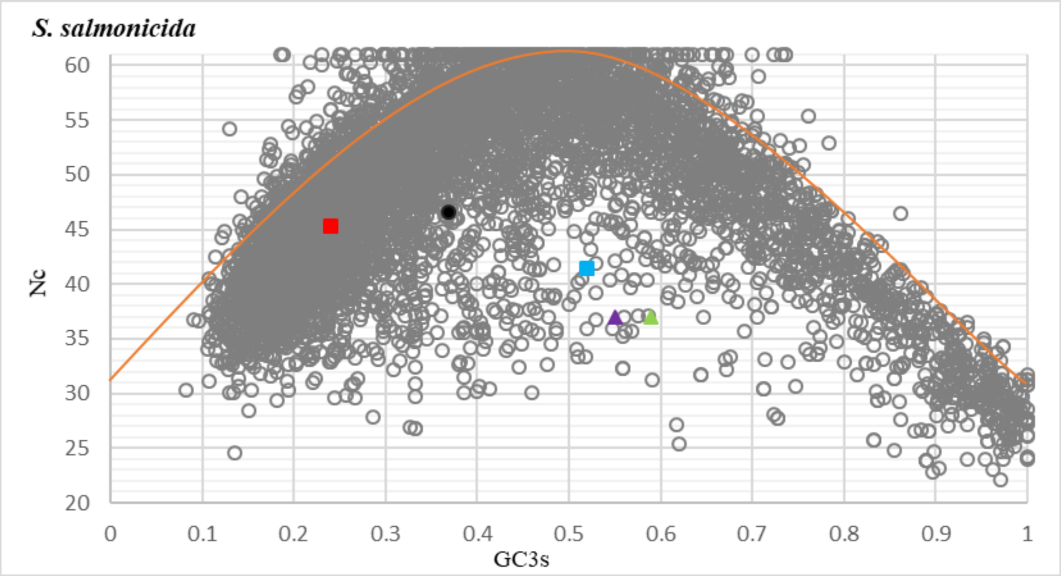


Figure 14 | Graph showing genome wide Nc plot for the Excavate *S. salmonicida*

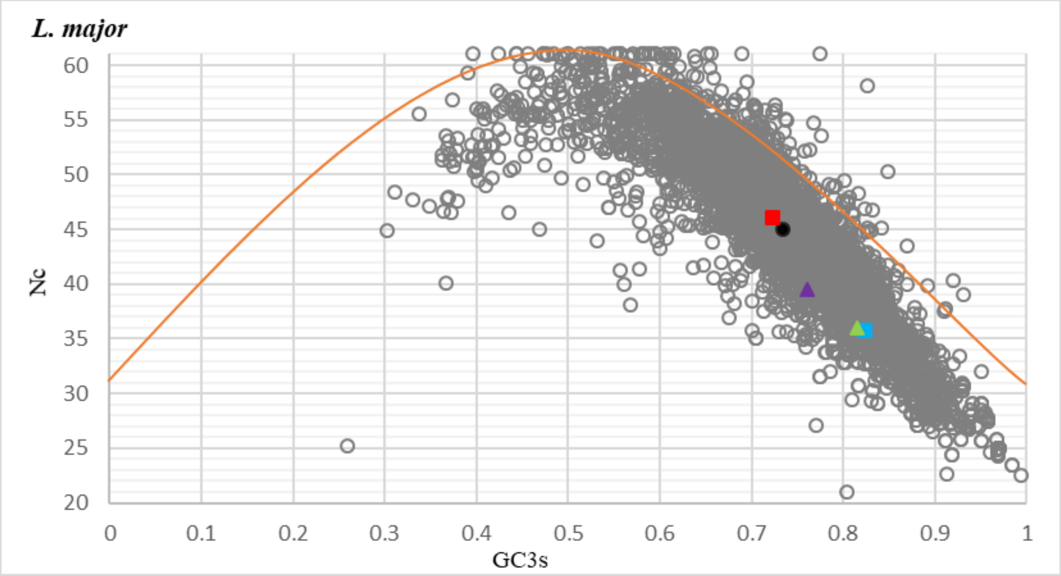


Figure 15 | Graph showing genome wide Nc plot for the Excavate *L. major*

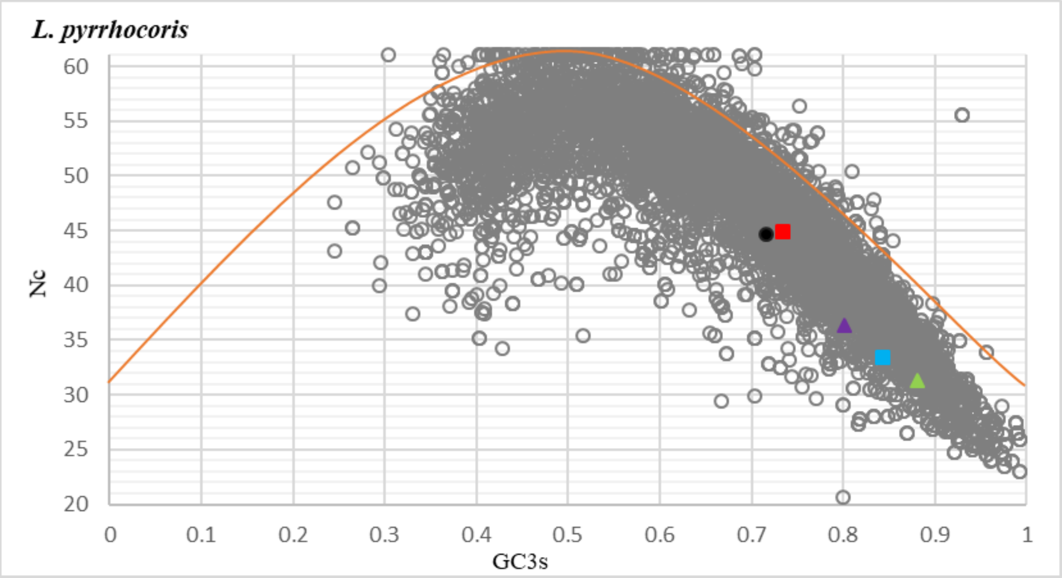


Figure 16 | Graph showing genome wide Nc plot for the Excavate *L. pyrrhocoris*

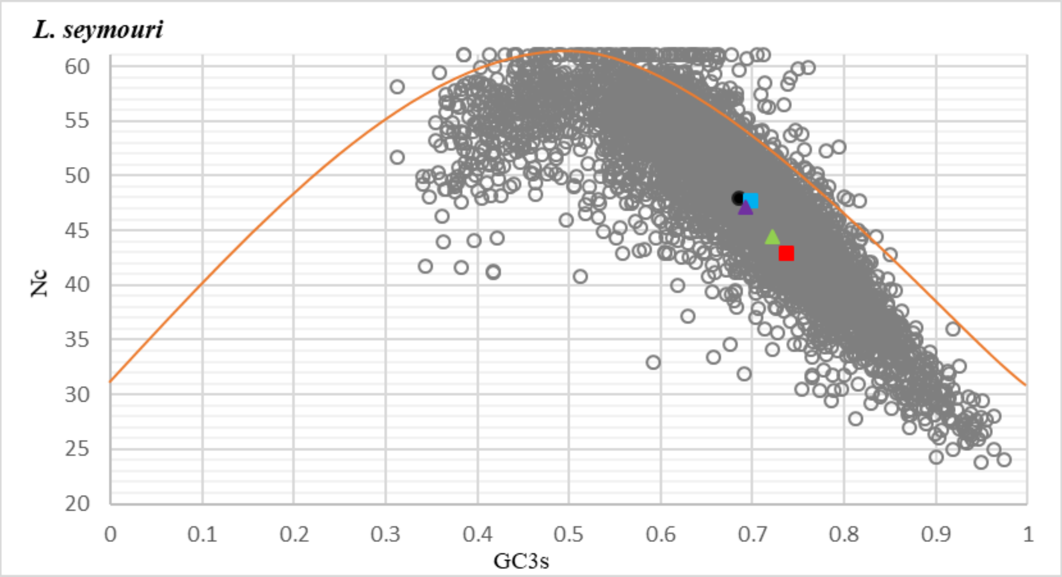


Figure 17 | Graph showing genome wide Nc plot for the Excavate *L. seymouri*

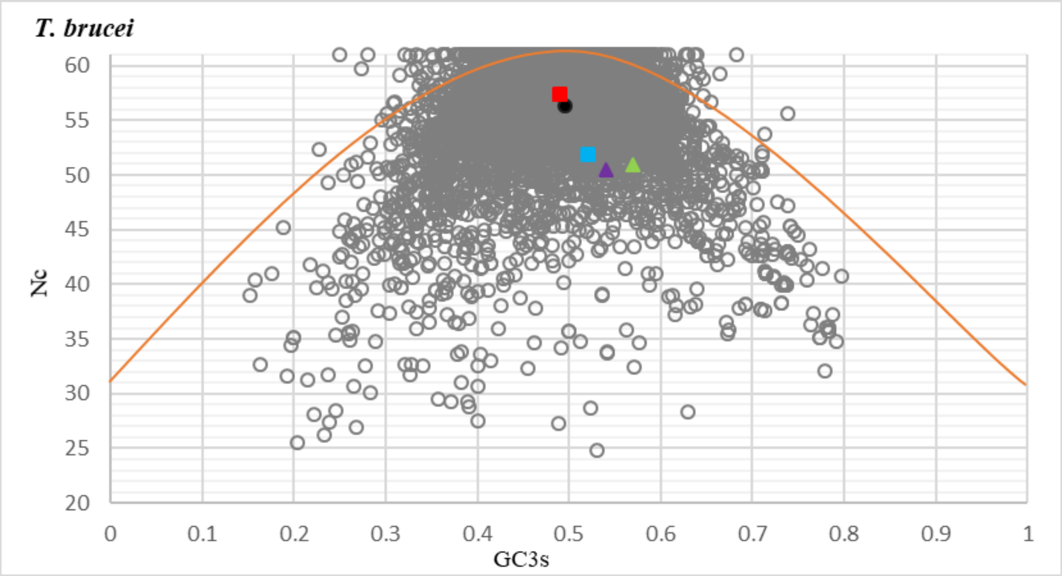


Figure 18 | Graph showing genome wide Nc plot for the Excavate *T. brucei*

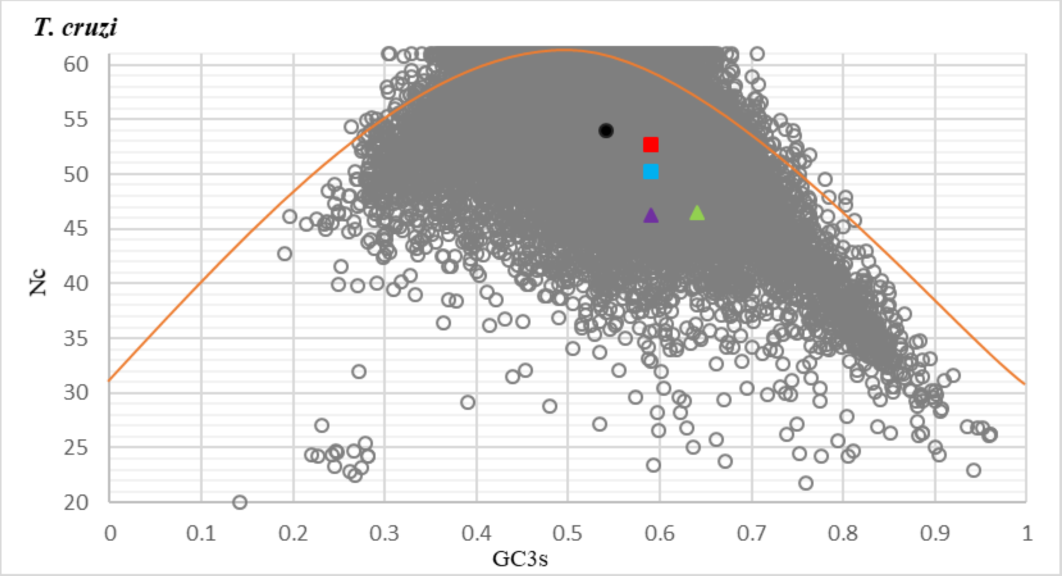


Figure 19 | Graph showing genome wide Nc plot for the Excavate *T. cruzi*

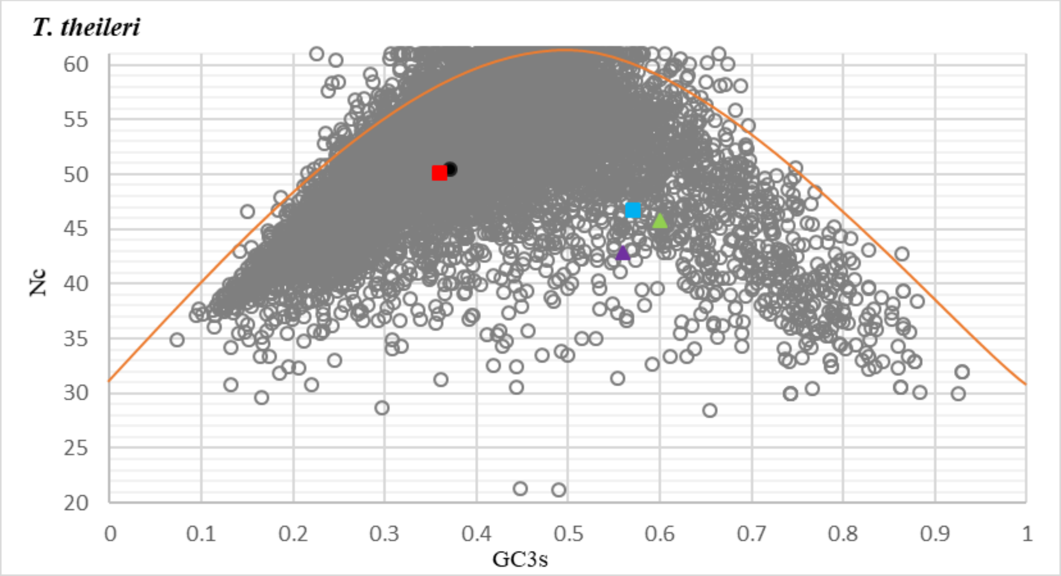


Figure 20 | Graph showing genome wide Nc plot for the Excavate *T. theileri*

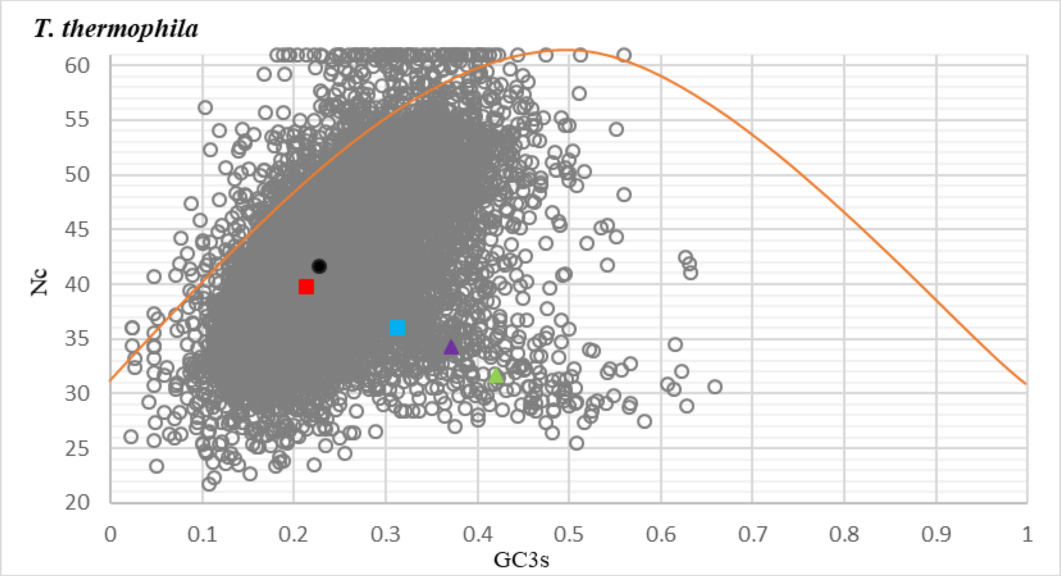


Figure 21 | Graph showing genome wide Nc plot for the Alveolate *T. thermophila*

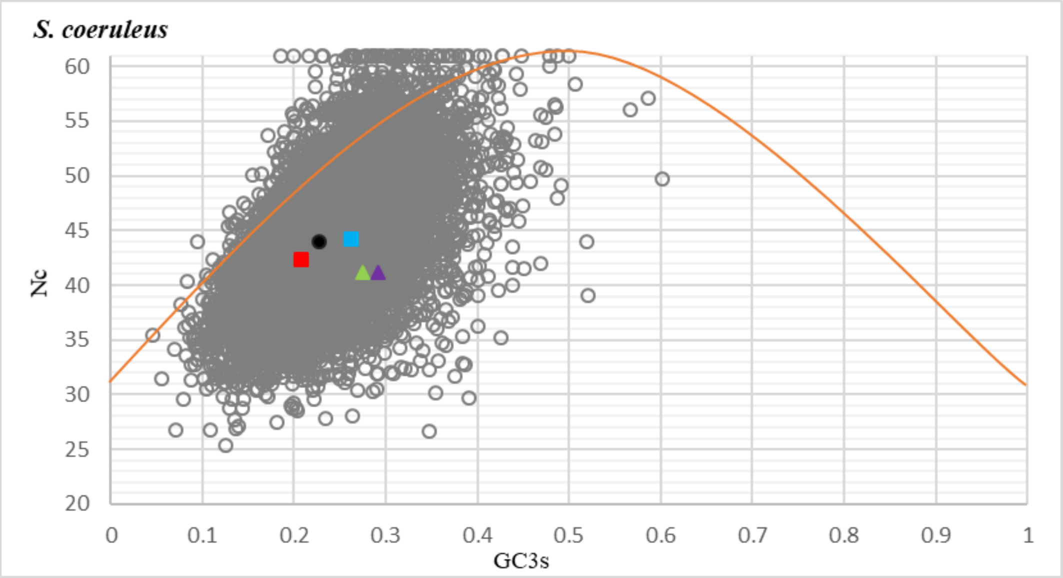


Figure 22 | Graph showing genome wide Nc plot for the Alveolate *S. coeruleus*

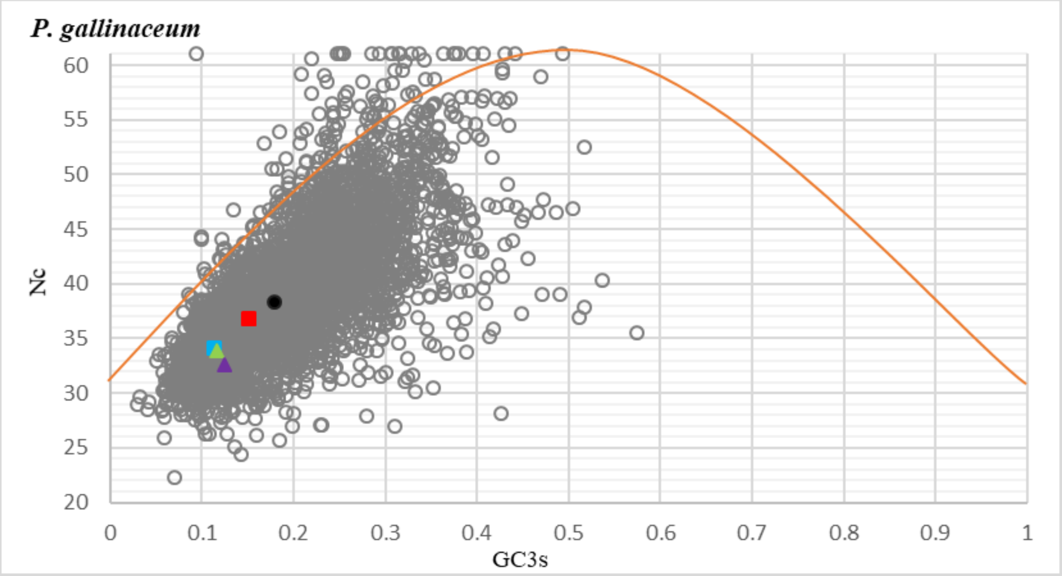


Figure 23 | Graph showing genome wide Nc plot for the Alveolate *P. gallinaceum*

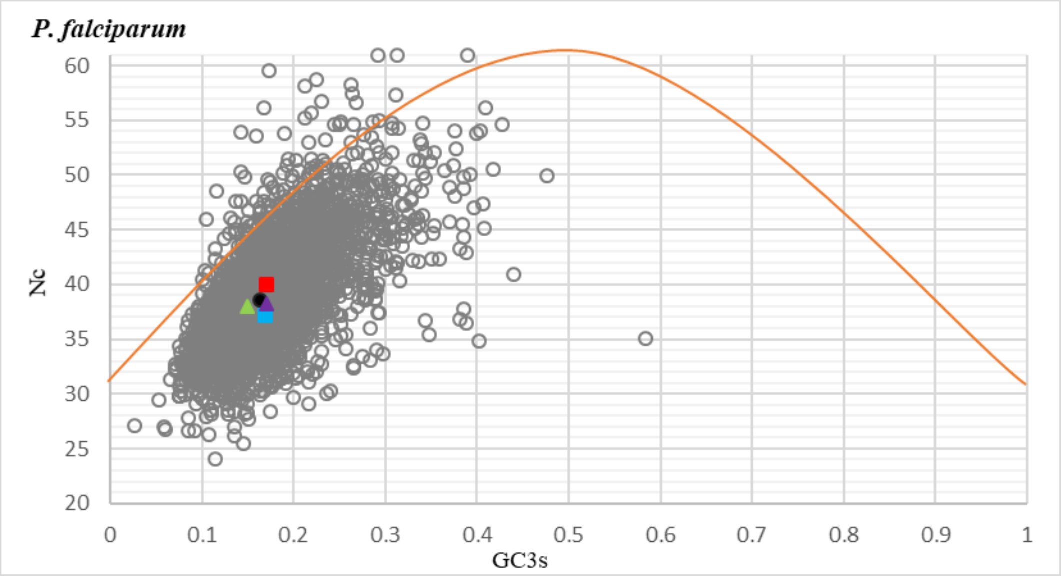


Figure 24 | Graph showing genome wide Nc plot for the Alveolate *P. falciparum*

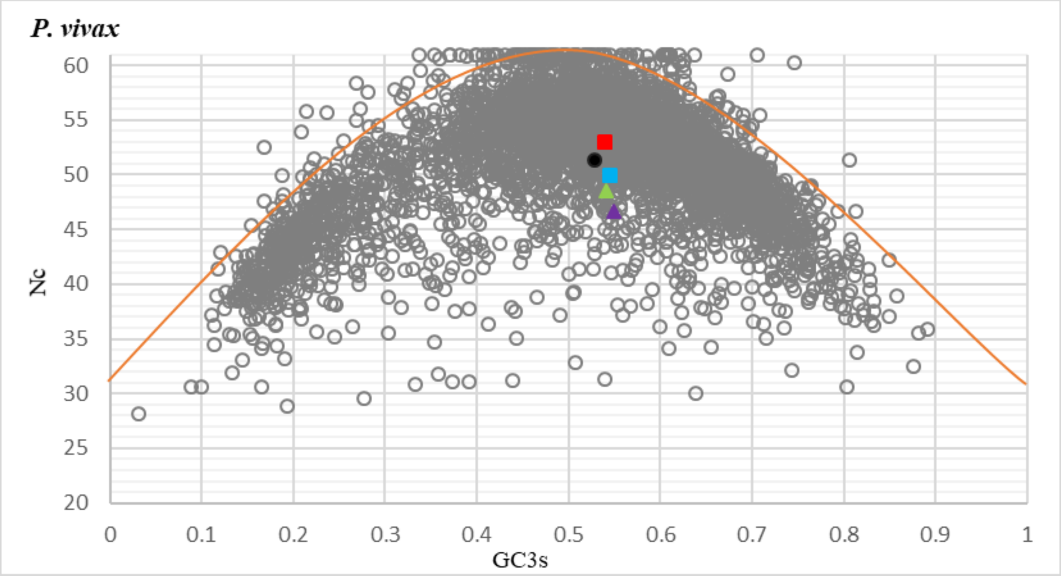


Figure 25 | Graph showing genome wide Nc plot for the Alveolate *P. vivax*

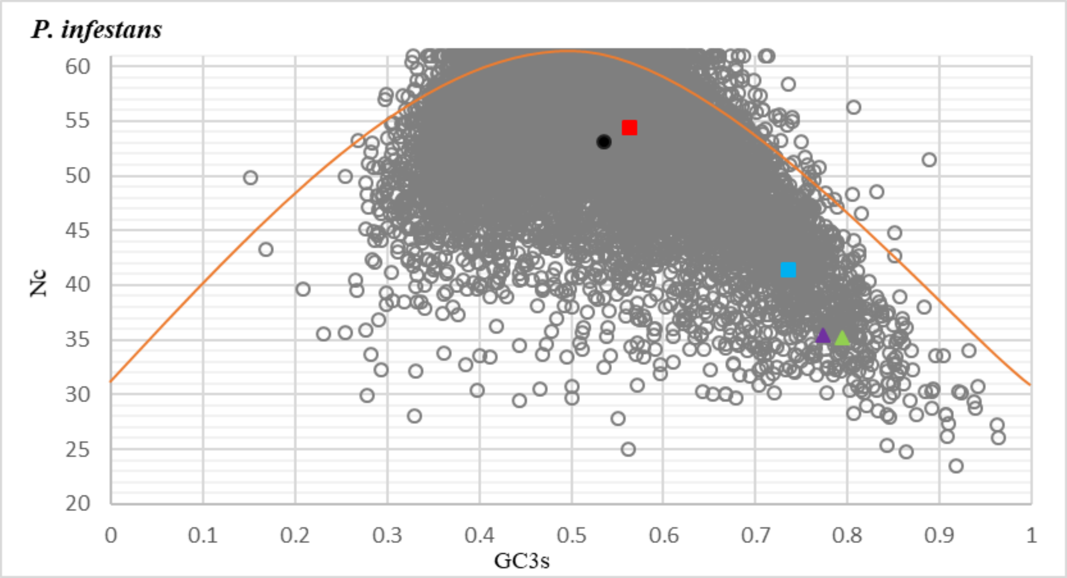


Figure 26 | Graph showing genome wide Nc plot for the Stramenopile *P. infestans*

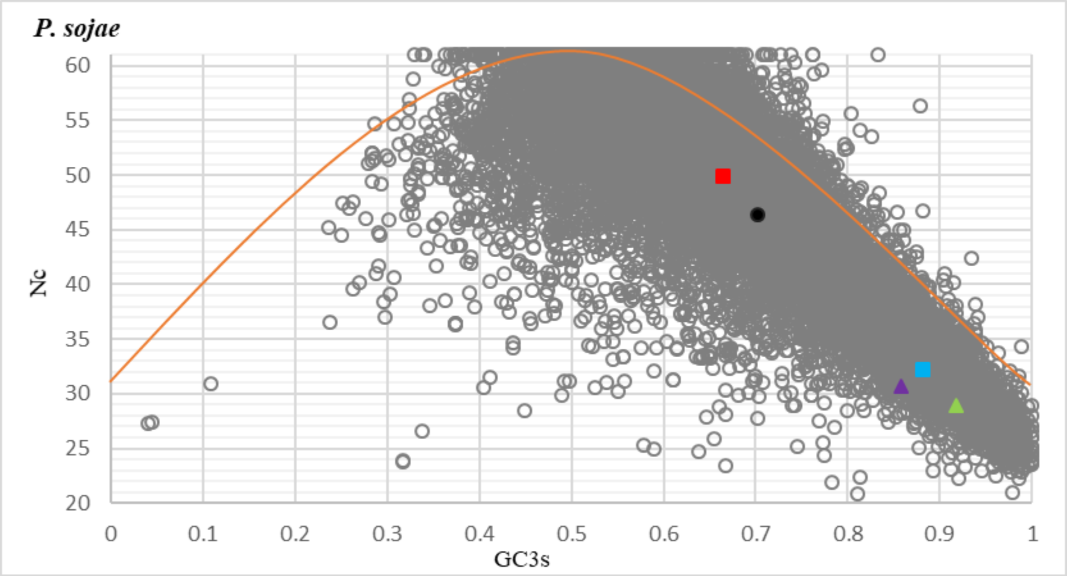


Figure 27 | Graph showing genome wide Nc plot for the Stramenopile *P. sojae*

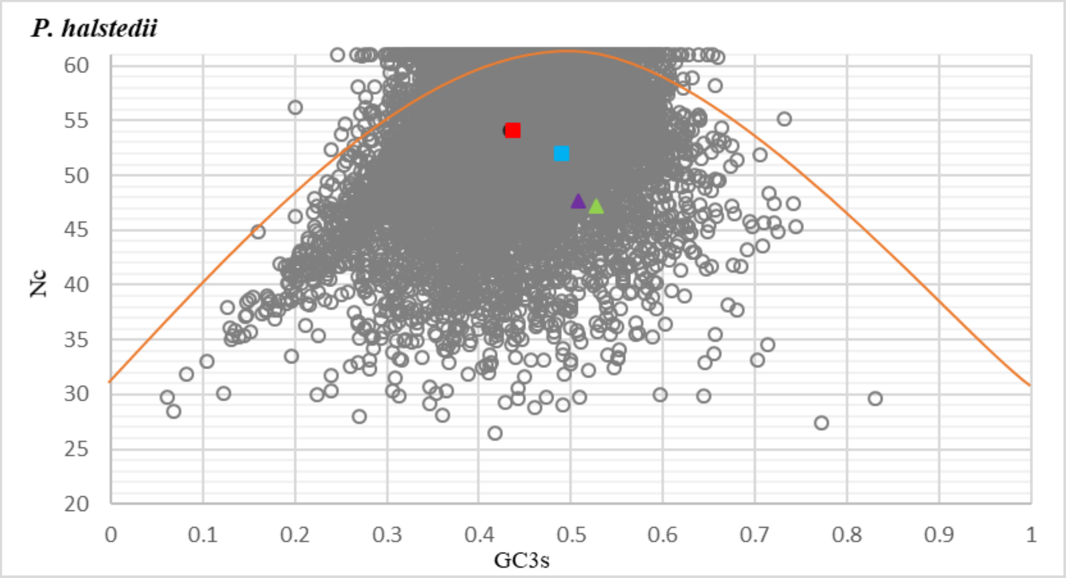


Figure 28 | Graph showing genome wide Nc plot for the Stramenopile *P. halstedii*

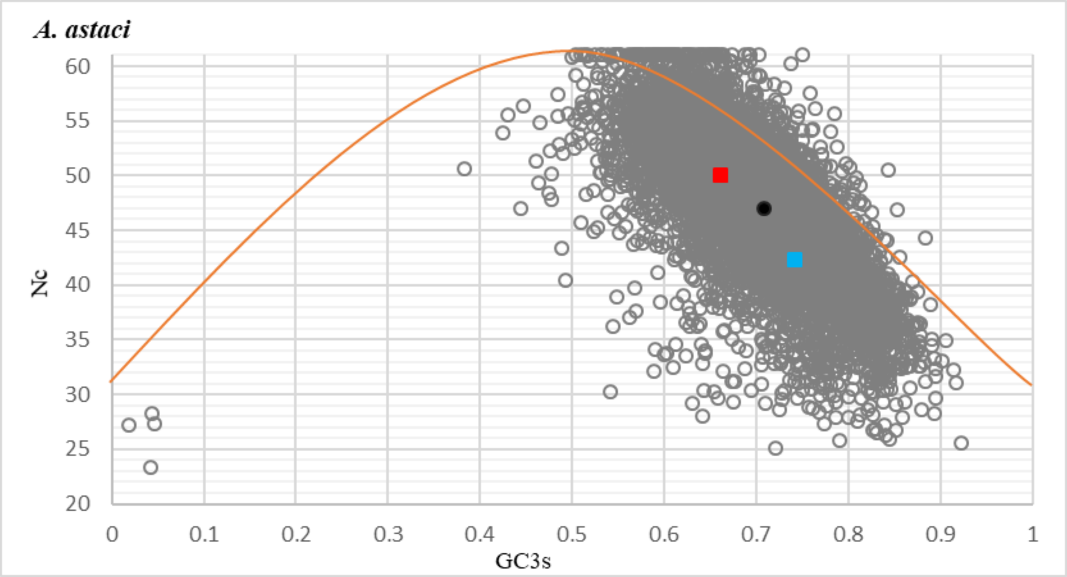


Figure 29 | Graph showing genome wide Nc plot for the Stramenopile *A. astaci*

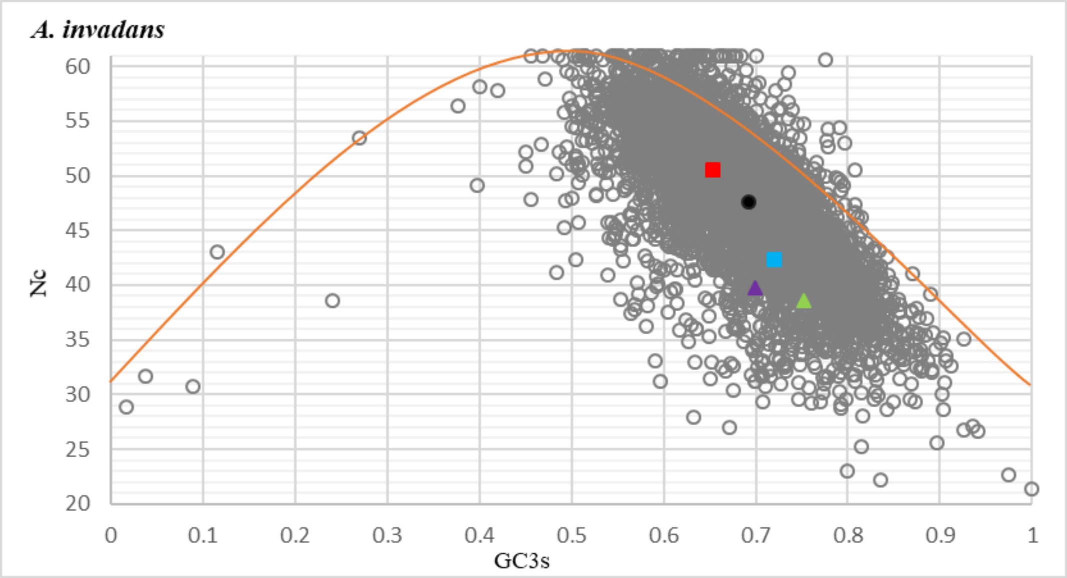


Figure 30 | Graph showing genome wide Nc plot for the Stramenopile *A. invadans*

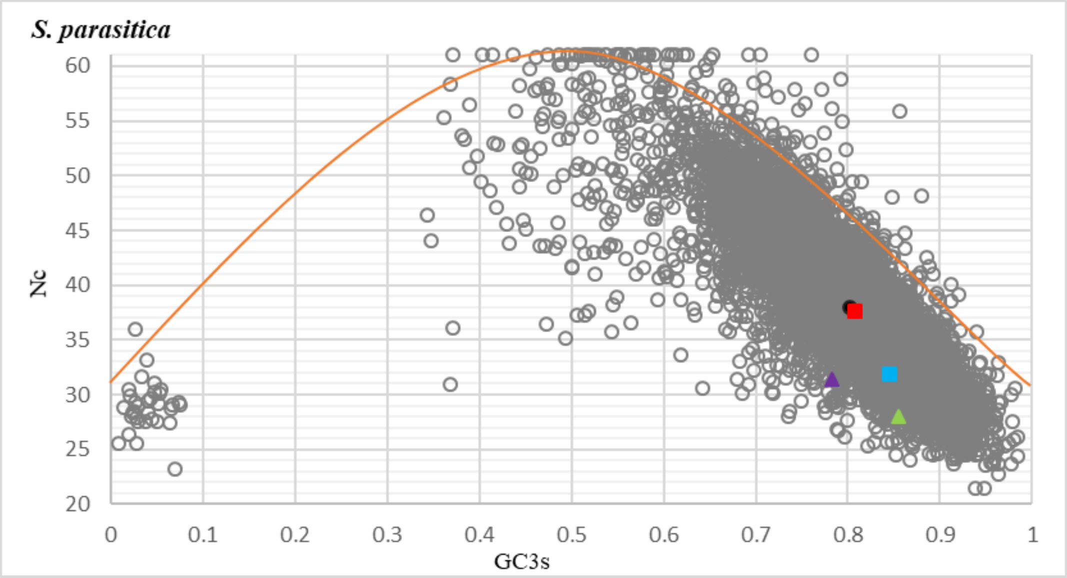


Figure 31 | Graph showing genome wide Nc plot for the Stramenopile *S. parasitica*

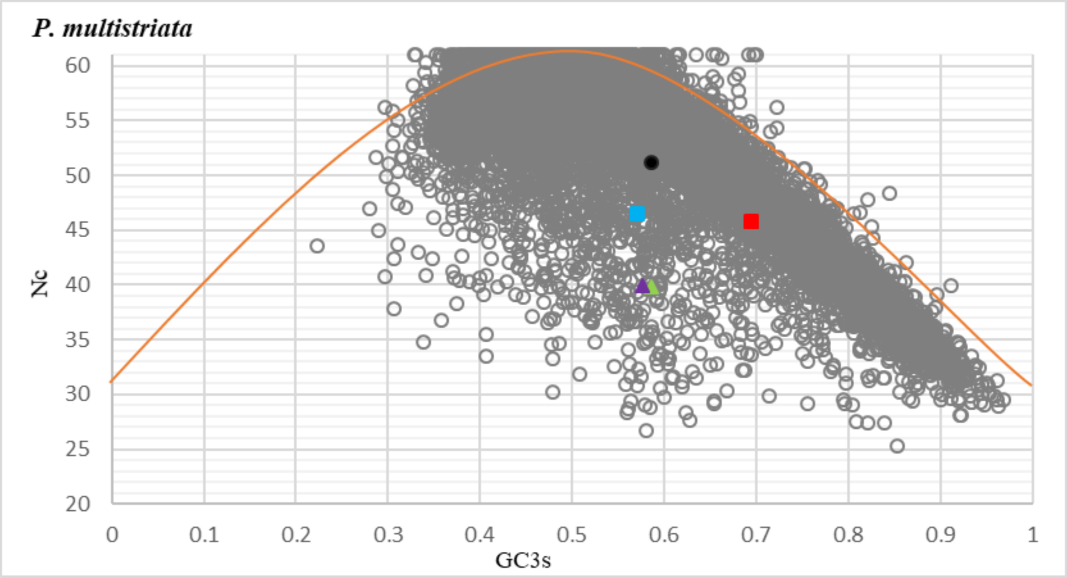


Figure 32 | Graph showing genome wide Nc plot for the Stramenopile *P. multistriata*

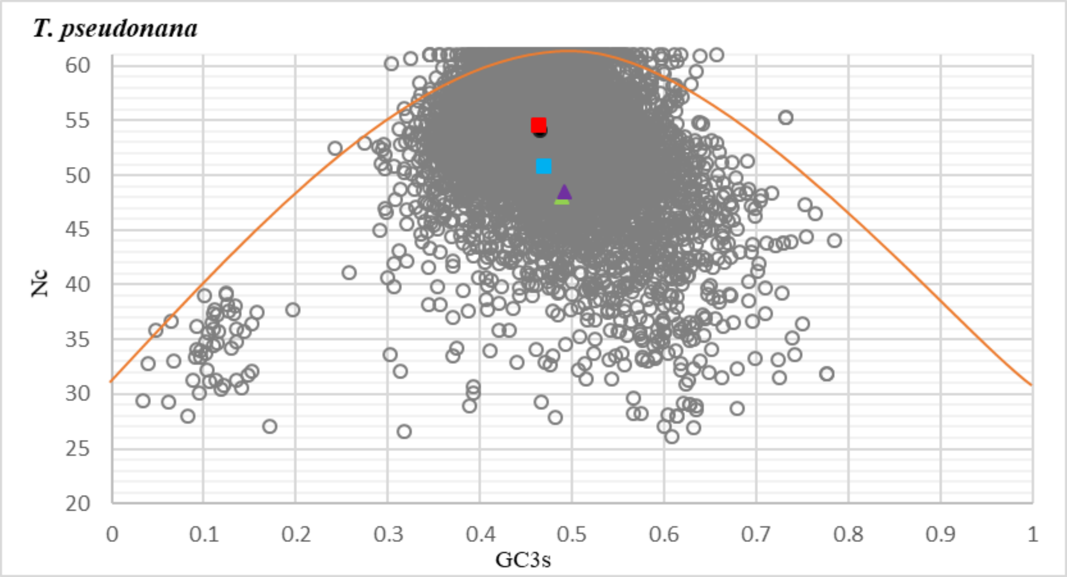


Figure 33 | Graph showing genome wide Nc plot for the Stramenopile *T. pseudonana*

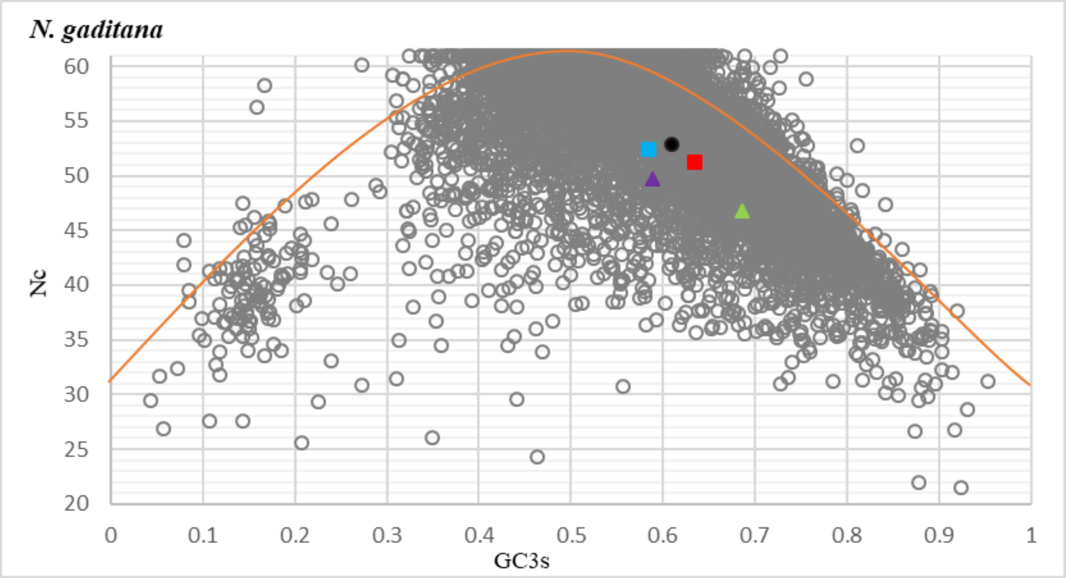


Figure 34 | Graph showing genome wide Nc plot for the Stramenopile *N. gaditana*

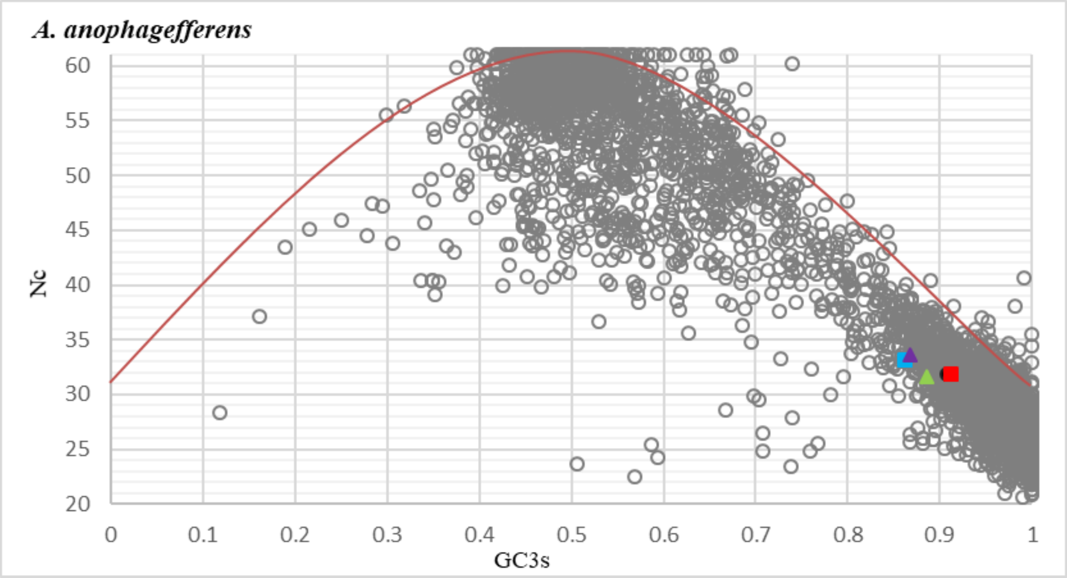


Figure 35 | Graph showing genome wide Nc plot for the Stramenopile *A. anophagefferens*

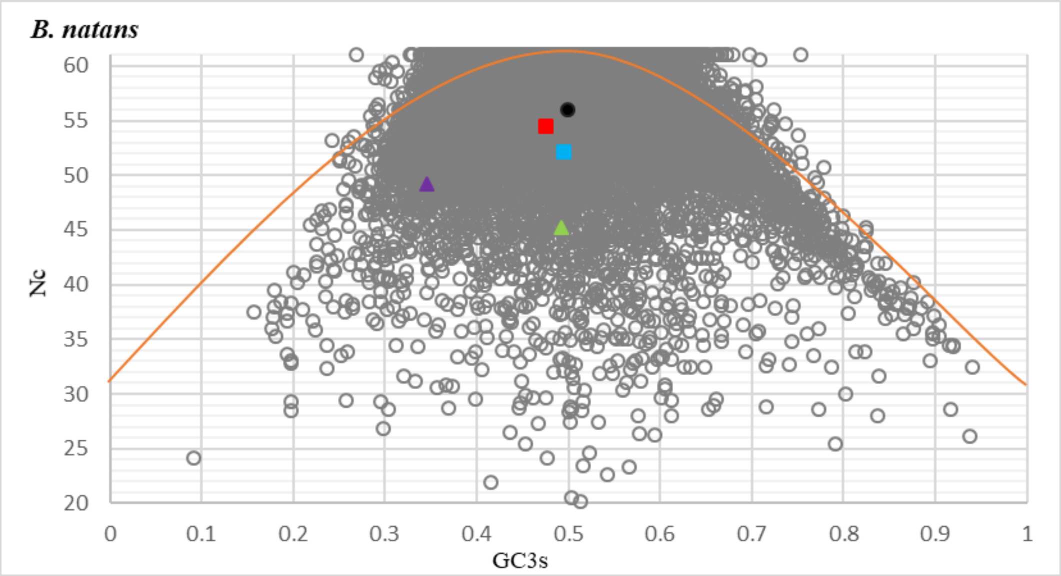


Figure 36 | Graph showing genome wide Nc plot for the Rhizarian *B. natans*

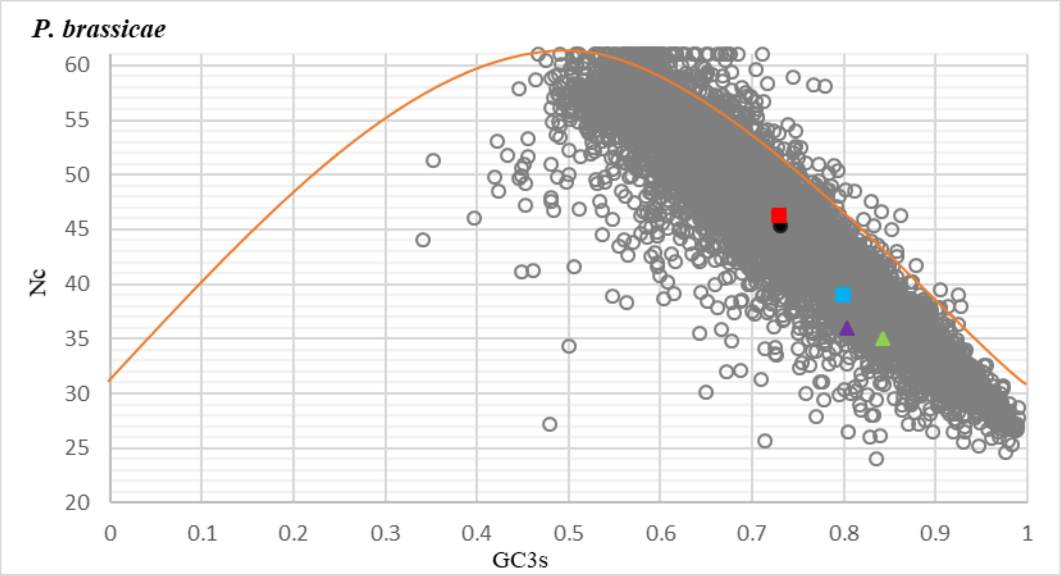


Figure 37 | Graph showing genome wide Nc plot for the Rhizarian *P. brassicae*

6.2 Codon Usage – Correspondence Analysis

General description of tables 12-42: The codon usage data is given in two forms, inside and outside of brackets. The numbers inside of the brackets represent the amount the codon was found within the genome files. The numbers outside the brackets represent the relative synonymous codon usage (RSCU) (Sharp, 1986).

Amino Acids	Codons	Highly Biased Genes	Weakly Biased Genes	Amino Acids	Codons	Highly Biased Genes	Weakly Biased Genes
Phe	UUU	0.18 (988)	0.72 (922)	**Ser**	UCU	0.24 (652)	0.80 (7033)
	UUC*	1.82 (10211)	1.28 (1632)		UCC*	1.58 (4298)	0.83 (7333)
Leu	UUA	0.00 (13)	0.24 (629)		UCA	0.09 (257)	1.37 (12096)
	UUG	0.26 (1126)	0.82 (2176)		UCG*	2.27 (6190)	1.33 (11740)
	CUU	0.21 (907)	1.39 (3689)		AGU	0.09 (239)	0.57 (4994)
	CUC*	3.21 (13960)	1.34 (3555)		AGC*	1.73 (4733)	1.11 (9833)
	CUA	0.03 (151)	1.04 (2749)	**Thr**	ACU	0.22 (843)	0.68 (5417)
	CUG*	2.28 (9899)	1.17 (3111)		ACC*	2.57 (9719)	0.86 (6867)
Val	GUU	0.15 (743)	1.03 (2387)		ACA	0.09 (337)	1.16 (9243)
	GUC*	1.78 (8706)	1.12 (2575)		ACG	1.12 (4215)	1.29 (10251)
	GUA	0.06 (277)	0.72 (1663)	**Tyr**	UAU	0.15 (632)	0.94 (450)
	GUG*	2.02 (9890)	1.13 (2608)		UAC*	1.85 (7717)	1.06 (509)
Pro	CCU	0.24 (695)	0.92 (9802)	**Ala**	GCU	0.29 (1939)	0.91 (8474)
	CCC*	2.32 (6639)	0.59 (6227)		GCC*	2.71 (18086)	0.70 (6469)
	CCA	0.09 (256)	1.32 (14028)		GCA	0.08 (546)	1.16 (10787)
	CCG*	1.34 (3845)	1.17 (12402)		GCG	0.92 (6168)	1.22 (11331)
His	CAU	0.12 (374)	1.15 (5704)	**Ile**	AUU	0.28 (1219)	0.93 (975)
	CAC*	1.88 (5681)	0.85 (4181)		AUC*	2.69 (11717)	1.55 (1628)
Gln	CAA	0.13 (759)	1.53 (7903)		AUA	0.03 (139)	0.52 (549)
	CAG*	1.87 (10905)	0.47 (2405)	**Asp**	GAU	0.31 (2557)	1.06 (2928)
Asn	AAU	0.13 (671)	0.93 (947)		GAC*	1.69 (13951)	0.94 (2614)
	AAC*	1.87 (9991)	1.07 (1086)	**Glu**	GAA	0.20 (2219)	1.30 (4147)
Lys	AAA	0.06 (583)	0.85 (1491)		GAG*	1.80 (20096)	0.70 (2249)
	AAG*	1.94 (20245)	1.15 (2022)	**Gly**	GGU	0.53 (2327)	0.74 (4400)
Cys	UGU	0.15 (205)	0.68 (4267)		GGC*	2.98 (13067)	1.09 (6524)
	UGC*	1.85 (2602)	1.32 (8262)		GGA	0.24 (1033)	1.27 (7598)
Arg	CGU	0.43 (1003)	0.66 (7416)		GGG	0.26 (1133)	0.90 (5350)
	CGC*	3.29 (7761)	1.00 (11174)	**Met**	AUG	1.00 (6214)	1.00 (2173)
	CGA	0.41 (959)	1.12 (12618)	**Trp**	UGG	1.00 (3535)	1.00 (9481)
	CGG	0.91 (2154)	0.89 (10000)	**TER**	UAA	1.38 (329)	0.19 (446)
	AGA	0.09 (216)	1.06 (11903)		UAG	0.62 (148)	0.16 (367)
	AGG	0.87 (2058)	1.27 (14205)		UGA	1.00 (239)	2.65 (6091)

Table 12A | Correspondence Analysis Results for the Amoebozoan *A. castellani*

Amino Acids	Codons	Highly Biased Genes	Weakly Biased Genes	Amino Acids	Codons	Highly Biased Genes	Weakly Biased Genes
Phe	UUU*	1.77 (9798)	0.99 (1453)	**Ser**	UCU*	1.75 (4739)	1.55 (1206)
	UUC	0.23 (1270)	1.01 (1479)		UCC	0.09 (251)	0.30 (233)
Leu	UUA*	3.96 (12071)	2.01 (2263)		UCA	2.81 (7598)	2.76 (2152)
	UUG	0.33 (1019)	0.48 (541)		UCG*	0.08 (229)	0.03 (26)
	CUU	1.29 (3924)	3.06 (3439)		AGU*	1.22 (3315)	1.11 (866)
	CUC	0.13 (394)	0.37 (417)		AGC	0.04 (114)	0.25 (193)
	CUA*	0.26 (804)	0.06 (66)	**Thr**	ACU	1.43 (3592)	2.36 (2935)
	CUG*	0.03 (96)	0.02 (18)		ACC	0.10 (251)	0.48 (595)
Val	GUU	2.19 (3811)	2.86 (4817)		ACA*	2.36 (5912)	1.10 (1367)
	GUC	0.18 (321)	0.54 (909)		ACG*	0.10 (259)	0.07 (82)
	GUA*	1.42 (2472)	0.55 (926)	**Tyr**	UAU*	1.87 (7797)	1.50 (1831)
	GUG*	0.21 (371)	0.05 (80)		UAC	0.13 (526)	0.50 (617)
Pro	CCU*	1.40 (2039)	0.34 (298)	**Ala**	GCU	1.44 (948)	2.56 (5699)
	CCC	0.06 (89)	0.04 (32)		GCC	0.09 (60)	0.36 (810)
	CCA	2.50 (3653)	3.60 (3126)		GCA*	2.37 (1555)	1.06 (2353)
	CCG	0.04 (57)	0.02 (18)		GCG*	0.10 (63)	0.02 (39)
His	CAU*	1.88 (3034)	1.46 (1326)	**Ile**	AUU	1.87 (13544)	2.47 (4585)
	CAC	0.12 (195)	0.54 (487)		AUC	0.14 (978)	0.38 (709)
Gln	CAA	1.91 (5957)	1.95 (3555)		AUA*	0.99 (7159)	0.14 (267)
	CAG*	0.09 (286)	0.05 (98)	**Asp**	GAU*	1.72 (5479)	1.62 (3902)
Asn	AAU*	1.83 (12309)	1.45 (3071)		GAC	0.28 (889)	0.38 (926)
	AAC	0.17 (1113)	0.55 (1179)	**Glu**	GAA*	1.86 (8991)	1.80 (7753)
Lys	AAA*	1.80 (10562)	1.28 (8657)		GAG	0.14 (680)	0.20 (876)
	AAG	0.20 (1179)	0.72 (4899)	**Gly**	GGU	0.97 (1450)	1.25 (1827)
Cys	UGU*	1.93 (6279)	1.35 (970)		GGC	0.05 (74)	0.05 (70)
	UGC	0.07 (219)	0.65 (470)		GGA	2.71 (4053)	2.64 (3851)
Arg	CGU*	0.92 (579)	0.57 (486)		GGG*	0.27 (404)	0.05 (76)
	CGC	0.04 (24)	0.03 (23)	**Met**	AUG	1.00 (3154)	1.00 (2550)
	CGA*	0.68 (423)	0.17 (145)	**Trp**	UGG	1.00 (1214)	1.00 (590)
	CGG*	0.06 (37)	0.01 (12)	**TER**	UAA	2.20 (293)	2.72 (362)
	AGA	4.03 (2523)	5.14 (4412)		UAG	0.25 (33)	0.05 (7)
	AGG*	0.27 (171)	0.08 (72)		UGA	0.55 (73)	0.23 (30)

Table 13A | Correspondence Analysis Results for the Amoebozoan *E. histolytica*

Amino Acids	Codons	Highly Biased Genes	Weakly Biased Genes	Amino Acids	Codons	Highly Biased Genes	Weakly Biased Genes
Phe	UUU*	1.68 (12413)	1.14 (9990)	**Ser**	UCU	0.78 (4001)	0.93 (3512)
	UUC	0.32 (2399)	0.86 (7524)		UCC	0.11 (560)	0.81 (3042)
Leu	UUA*	4.95 (22132)	1.59 (9451)		UCA*	3.45 (17732)	2.10 (7908)
	UUG	0.37 (1640)	1.46 (8702)		UCG	0.08 (400)	0.64 (2418)
	CUU	0.41 (1830)	0.78 (4646)		AGU*	1.51 (7763)	0.77 (2899)
	CUC	0.04 (194)	0.67 (3980)		AGC	0.07 (385)	0.74 (2798)
	CUA	0.22 (990)	0.81 (4817)	**Thr**	ACU*	1.37 (5183)	0.78 (3016)
	CUG	0.01 (26)	0.69 (4095)		ACC	0.28 (1051)	1.20 (4615)
Val	GUU*	2.49 (5603)	0.99 (4520)		ACA*	2.30 (8676)	1.66 (6412)
	GUC	0.12 (259)	0.53 (2439)		ACG	0.06 (209)	0.35 (1363)
	GUA	1.28 (2887)	1.36 (6231)	**Tyr**	UAU*	1.88 (10101)	1.15 (8539)
	GUG	0.11 (252)	1.11 (5091)		UAC	0.12 (658)	0.85 (6264)
Pro	CCU	0.59 (1331)	0.91 (1765)	**Ala**	GCU*	1.18 (1465)	0.77 (848)
	CCC	0.07 (152)	0.55 (1066)		GCC	0.16 (205)	0.93 (1025)
	CCA*	3.30 (7388)	2.03 (3922)		GCA*	2.60 (3237)	1.91 (2097)
	CCG	0.04 (86)	0.50 (973)		GCG	0.06 (78)	0.39 (433)
His	CAU*	1.89 (3557)	1.14 (8063)	**Ile**	AUU*	1.99 (18939)	0.95 (13323)
	CAC	0.11 (208)	0.86 (6138)		AUC	0.16 (1514)	0.73 (10264)
Gln	CAA*	1.96 (12075)	1.52 (15498)		AUA	0.85 (8037)	1.31 (18330)
	CAG	0.04 (231)	0.48 (4865)	**Asp**	GAU*	1.92 (14733)	1.27 (4107)
Asn	AAU*	1.93 (55930)	1.01 (13684)		GAC	0.08 (649)	0.73 (2377)
	AAC	0.07 (1960)	0.99 (13284)	**Glu**	GAA*	1.82 (15137)	1.46 (7048)
Lys	AAA*	1.83 (21783)	1.32 (20709)		GAG	0.18 (1539)	0.54 (2588)
	AAG	0.17 (2001)	0.68 (10665)	**Gly**	GGU*	2.97 (8270)	1.13 (1814)
Cys	UGU*	1.91 (3653)	1.06 (3780)		GGC	0.09 (260)	0.64 (1024)
	UGC	0.09 (173)	0.94 (3378)		GGA	0.88 (2442)	1.46 (2335)
Arg	CGU*	0.73 (715)	0.30 (799)		GGG	0.06 (162)	0.77 (1237)
	CGC	0.01 (11)	0.23 (615)	**Met**	AUG	1.00 (3694)	1.00 (9994)
	CGA	0.09 (86)	0.59 (1540)	**Trp**	UGG	1.00 (1742)	1.00 (6861)
	CGG	0.01 (9)	0.19 (512)	**TER**	UAA	2.68 (584)	1.46 (11949)
	AGA*	4.84 (4720)	3.45 (9079)		UAG	0.16 (35)	0.68 (5526)
	AGG	0.32 (310)	1.23 (3248)		UGA	0.16 (34)	0.86 (7010)

Table 14A | Correspondence Analysis Results for the Amoebozoan *D. discoideum*

Amino Acids	Codons	Highly Biased Genes	Weakly Biased Genes	Amino Acids	Codons	Highly Biased Genes	Weakly Biased Genes
Phe	UUU	0.46 (4027)	0.88 (6746)	Ser	UCU*	1.58 (9419)	1.35 (15509)
	UUC*	1.54 (13411)	1.12 (8575)		UCC*	1.67 (9942)	0.81 (9312)
Leu	UUA	0.31 (1764)	0.51 (3936)		UCA	0.47 (2823)	1.55 (17834)
	UUG*	1.08 (6152)	0.62 (4742)		UCG	0.66 (3900)	0.91 (10505)
	CUU	0.42 (2410)	1.69 (13046)		AGU	0.40 (2380)	0.69 (7893)
	CUC*	2.61 (14867)	1.51 (11664)		AGC*	1.21 (7202)	0.69 (7928)
	CUA	0.18 (1052)	0.74 (5735)	Thr	ACU*	0.89 (7240)	0.82 (7600)
	CUG*	1.40 (7953)	0.93 (7135)		ACC*	1.93 (15805)	0.87 (8133)
Val	GUU	0.57 (4287)	1.20 (7865)		ACA	0.55 (4518)	1.53 (14299)
	GUC*	2.23 (16712)	1.11 (7281)		ACG	0.63 (5142)	0.78 (7251)
	GUA	0.34 (2528)	0.77 (5024)	Tyr	UAU	0.31 (2176)	1.01 (5062)
	GUG	0.86 (6419)	0.91 (5962)		UAC*	1.69 (12074)	0.99 (4919)
Pro	CCU	0.83 (4036)	1.12 (10113)	Ala	GCU	1.00 (9435)	1.17 (6546)
	CCC*	2.31 (11206)	0.66 (6015)		GCC*	2.18 (20499)	0.79 (4460)
	CCA	0.61 (2933)	1.45 (13110)		GCA	0.54 (5068)	1.51 (8463)
	CCG	0.25 (1197)	0.77 (7024)		GCG	0.28 (2621)	0.53 (3000)
His	CAU	0.31 (1095)	1.18 (19962)	Ile	AUU	0.42 (3058)	1.15 (9444)
	CAC*	1.69 (5947)	0.82 (13966)		AUC*	2.49 (18325)	1.33 (10940)
Gln	CAA	0.64 (5619)	1.26 (19445)		AUA	0.09 (699)	0.52 (4265)
	CAG*	1.36 (11978)	0.74 (11487)	Asp	GAU	0.78 (8392)	1.12 (12556)
Asn	AAU	0.31 (3747)	1.03 (5882)		GAC*	1.22 (13012)	0.88 (9793)
	AAC*	1.69 (20542)	0.97 (5526)	Glu	GAA	0.51 (5412)	1.16 (18981)
Lys	AAA	0.28 (3337)	1.05 (7811)		GAG*	1.49 (15708)	0.84 (13736)
	AAG*	1.72 (20897)	0.95 (7028)	Gly	GGU*	1.29 (10557)	0.72 (6724)
Cys	UGU	0.68 (2074)	1.05 (9416)		GGC	0.80 (6501)	0.88 (8219)
	UGC*	1.32 (4055)	0.95 (8511)		GGA*	1.81 (14760)	1.52 (14247)
Arg	CGU*	1.03 (2678)	0.86 (11536)		GGG	0.10 (802)	0.88 (8191)
	CGC*	1.03 (2669)	0.59 (7898)	Met	AUG	1.00 (7467)	1.00 (5509)
	CGA	0.49 (1265)	1.36 (18357)	Trp	UGG	1.00 (5368)	1.00 (11868)
	CGG	0.08 (220)	0.75 (10079)	TER	UAA	1.75 (489)	0.60 (4314)
	AGA*	2.72 (7056)	1.73 (23263)		UAG	0.41 (115)	0.49 (3518)
	AGG	0.65 (1688)	0.72 (9753)		UGA	0.84 (233)	1.92 (13896)

Table 15A | Correspondence Analysis Results for the Amoebozoan *P. fungivorum*

Amino Acids	Codons	Highly Biased Genes	Weakly Biased Genes	Amino Acids	Codons	Highly Biased Genes	Weakly Biased Genes
Phe	UUU*	0.94 (4342)	0.86 (1616)	**Ser**	UCU	0.39 (1063)	0.90 (4106)
	UUC	1.06 (4904)	1.14 (2147)		UCC*	1.51 (4120)	0.91 (4139)
Leu	UUA	0.00 (6)	0.17 (619)		UCA	0.28 (774)	0.87 (3942)
	UUG	0.06 (226)	1.19 (4269)		UCG*	3.48 (9472)	1.54 (6998)
	CUU	0.64 (2275)	0.92 (3311)		AGU	0.04 (97)	0.56 (2557)
	CUC*	4.31 (15306)	1.58 (5648)		AGC	0.30 (826)	1.21 (5516)
	CUA	0.03 (111)	0.50 (1780)	**Thr**	ACU	0.42 (1317)	0.74 (2657)
	CUG	0.95 (3368)	1.64 (5884)		ACC*	2.75 (8576)	0.87 (3131)
Val	GUU	0.42 (1786)	0.73 (2679)		ACA	0.21 (662)	0.82 (2973)
	GUC*	2.86 (12216)	1.12 (4114)		ACG	0.62 (1920)	1.57 (5668)
	GUA	0.02 (88)	0.52 (1892)	**Tyr**	UAU	0.12 (391)	0.79 (822)
	GUG	0.70 (2981)	1.63 (5961)		UAC*	1.88 (6267)	1.21 (1269)
Pro	CCU	0.29 (781)	0.92 (5048)	**Ala**	GCU	0.57 (3919)	0.99 (7695)
	CCC*	1.34 (3578)	0.61 (3381)		GCC*	2.66 (18334)	0.85 (6577)
	CCA	0.18 (485)	0.99 (5428)		GCA	0.26 (1804)	0.77 (6024)
	CCG*	2.18 (5815)	1.48 (8180)		GCG	0.52 (3556)	1.39 (10802)
His	CAU	0.27 (622)	1.06 (4799)	**Ile**	AUU	0.53 (2006)	0.69 (756)
	CAC*	1.73 (4063)	0.94 (4240)		AUC*	2.46 (9285)	1.94 (2120)
Gln	CAA	0.09 (311)	1.28 (5578)		AUA	0.00 (16)	0.36 (395)
	CAG*	1.91 (6313)	0.72 (3114)	**Asp**	GAU	0.27 (1951)	0.83 (3195)
Asn	AAU	0.10 (380)	0.63 (581)		GAC*	1.73 (12344)	1.17 (4512)
	AAC*	1.90 (7384)	1.37 (1252)	**Glu**	GAA	0.14 (904)	0.64 (1930)
Lys	AAA	0.04 (238)	0.57 (803)		GAG*	1.86 (12409)	1.36 (4114)
	AAG*	1.96 (12146)	1.43 (2004)	**Gly**	GGU	0.56 (2135)	0.81 (4792)
Cys	UGU	0.16 (185)	0.73 (3695)		GGC*	3.14 (11901)	1.61 (9514)
	UGC*	1.84 (2189)	1.27 (6489)		GGA	0.12 (443)	0.77 (4574)
Arg	CGU	0.73 (1461)	0.81 (5818)		GGG	0.18 (665)	0.81 (4763)
	CGC*	4.44 (8920)	1.50 (10748)	**Met**	AUG	1.00 (5615)	1.00 (3269)
	CGA	0.07 (135)	1.28 (9153)	**Trp**	UGG	1.00 (2490)	1.00 (6165)
	CGG	0.69 (1384)	1.41 (10118)	**TER**	UAA	1.50 (265)	0.25 (304)
	AGA	0.02 (47)	0.39 (2809)		UAG	0.79 (140)	0.32 (392)
	AGG	0.05 (100)	0.60 (4285)		UGA	0.71 (126)	2.43 (2973)

Table 16A | Correspondence Analysis Results for the Apusomonadan *T. trahens*

Amino Acids	Codons	Highly Biased Genes	Weakly Biased Genes	Amino Acids	Codons	Highly Biased Genes	Weakly Biased Genes
Phe	UUU	0.08 (157)	0.99 (863)	**Ser**	UCU	0.48 (483)	1.03 (1512)
	UUC*	1.92 (3829)	1.01 (872)		UCC*	3.20 (3196)	1.26 (1845)
Leu	UUA	0.00 (0)	0.49 (678)		UCA	0.01 (10)	1.01 (1473)
	UUG	0.03 (33)	0.70 (965)		UCG	1.03 (1032)	1.33 (1942)
	CUU	0.34 (448)	1.31 (1807)		AGU	0.06 (57)	0.48 (709)
	CUC*	2.48 (3253)	1.76 (2420)		AGC*	1.22 (1218)	0.89 (1299)
	CUA	0.00 (6)	0.73 (1001)	**Thr**	ACU	0.35 (469)	0.78 (772)
	CUG*	3.14 (4116)	1.00 (1381)		ACC*	3.45 (4605)	1.10 (1085)
Val	GUU	0.43 (680)	0.74 (868)		ACA	0.01 (10)	1.19 (1168)
	GUC*	2.68 (4216)	1.39 (1627)		ACG	0.19 (259)	0.93 (911)
	GUA	0.00 (0)	0.66 (769)	**Tyr**	UAU	0.22 (294)	1.31 (635)
	GUG	0.89 (1400)	1.22 (1424)		UAC*	1.78 (2423)	0.69 (332)
Pro	CCU	0.17 (166)	1.02 (3367)	**Ala**	GCU	0.42 (934)	0.94 (2431)
	CCC*	2.29 (2271)	0.96 (3172)		GCC*	3.25 (7276)	1.15 (2971)
	CCA	0.01 (5)	0.92 (3018)		GCA	0.01 (18)	0.91 (2364)
	CCG*	1.54 (1523)	1.10 (3609)		GCG	0.32 (725)	1.01 (2612)
His	CAU	0.27 (268)	1.14 (2812)	**Ile**	AUU	0.42 (710)	1.03 (589)
	CAC*	1.73 (1725)	0.86 (2100)		AUC*	2.58 (4332)	1.18 (676)
Gln	CAA	0.02 (29)	1.33 (2450)		AUA	0.00 (1)	0.79 (456)
	CAG*	1.98 (3249)	0.67 (1224)	**Asp**	GAU	0.47 (1325)	1.06 (1492)
Asn	AAU	0.18 (360)	1.19 (556)		GAC*	1.53 (4290)	0.94 (1319)
	AAC*	1.82 (3749)	0.81 (382)	**Glu**	GAA	0.01 (46)	1.20 (1506)
Lys	AAA	0.01 (17)	1.15 (750)		GAG*	1.99 (6248)	0.80 (1005)
	AAG*	1.99 (6022)	0.85 (558)	**Gly**	GGU*	1.17 (1620)	0.72 (1909)
Cys	UGU	0.12 (66)	0.86 (1493)		GGC*	2.56 (3535)	1.44 (3829)
	UGC*	1.88 (1032)	1.14 (1994)		GGA	0.02 (23)	1.00 (2666)
Arg	CGU*	0.93 (865)	0.63 (1449)		GGG	0.25 (343)	0.84 (2243)
	CGC*	4.54 (4219)	1.16 (2660)	**Met**	AUG	1.00 (2310)	1.00 (1240)
	CGA	0.01 (9)	1.37 (3128)	**Trp**	UGG	1.00 (838)	1.00 (2309)
	CGG	0.51 (472)	1.60 (3675)	**TER**	UAA	2.33 (240)	0.43 (241)
	AGA	0.00 (0)	0.51 (1162)		UAG	0.23 (24)	0.30 (171)
	AGG	0.01 (11)	0.73 (1673)		UGA	0.44 (45)	2.27 (1285)

Table 17 | Correspondence Analysis Results for the Opisthokont *F. alba*

Amino Acids	Codons	Highly Biased Genes	Weakly Biased Genes	Amino Acids	Codons	Highly Biased Genes	Weakly Biased Genes
Phe	UUU	0.68 (2835)	0.91 (2540)	**Ser**	UCU	0.86 (1798)	0.87 (3739)
	UUC*	1.32 (5444)	1.09 (3051)		UCC*	1.57 (3310)	0.92 (3962)
Leu	UUA	0.00 (5)	0.29 (1000)		UCA	0.13 (279)	1.12 (4800)
	UUG	0.33 (958)	1.23 (4163)		UCG*	2.20 (4634)	1.19 (5093)
	CUU	0.82 (2351)	1.27 (4321)		AGU	0.06 (133)	0.65 (2785)
	CUC*	3.37 (9705)	1.30 (4430)		AGC	1.17 (2460)	1.25 (5380)
	CUA	0.01 (24)	0.54 (1846)	**Thr**	ACU	0.86 (2458)	0.82 (2773)
	CUG*	1.48 (4259)	1.36 (4622)		ACC*	2.39 (6861)	0.92 (3091)
Val	GUU	0.79 (2845)	1.06 (2846)		ACA	0.12 (338)	1.10 (3724)
	GUC*	2.60 (9374)	1.35 (3620)		ACG	0.64 (1835)	1.16 (3919)
	GUA	0.01 (32)	0.62 (1665)	**Tyr**	UAU	0.20 (577)	1.03 (990)
	GUG	0.60 (2182)	0.97 (2601)		UAC*	1.80 (5184)	0.97 (926)
Pro	CCU	0.45 (934)	0.84 (4041)	**Ala**	GCU	1.00 (5382)	1.10 (6503)
	CCC*	2.55 (5303)	0.68 (3282)		GCC*	2.35 (12637)	0.76 (4491)
	CCA	0.19 (397)	1.37 (6650)		GCA	0.22 (1182)	1.14 (6736)
	CCG	0.82 (1698)	1.11 (5381)		GCG	0.42 (2267)	1.01 (5976)
His	CAU	0.15 (326)	1.06 (5643)	**Ile**	AUU*	1.27 (4603)	1.02 (1551)
	CAC*	1.85 (3948)	0.94 (4981)		AUC*	1.72 (6237)	1.56 (2362)
Gln	CAA	0.45 (1673)	1.36 (8119)		AUA	0.00 (12)	0.42 (642)
	CAG*	1.55 (5699)	0.64 (3821)	**Asp**	GAU	0.38 (2217)	0.90 (2844)
Asn	AAU	0.08 (341)	0.88 (1742)		GAC*	1.62 (9430)	1.10 (3461)
	AAC*	1.92 (7999)	1.12 (2207)	**Glu**	GAA	0.29 (1873)	1.19 (3056)
Lys	AAA	0.04 (287)	1.11 (2869)		GAG*	1.71 (10928)	0.81 (2099)
	AAG*	1.96 (13750)	0.89 (2291)	**Gly**	GGU	0.76 (2597)	0.78 (2662)
Cys	UGU	0.06 (63)	0.71 (4265)		GGC*	3.12 (10698)	1.30 (4394)
	UGC*	1.94 (2136)	1.29 (7796)		GGA	0.10 (330)	1.22 (4148)
Arg	CGU*	1.13 (1962)	0.94 (5369)		GGG	0.02 (84)	0.70 (2364)
	CGC*	4.49 (7776)	1.30 (7432)	**Met**	AUG	1.00 (4163)	1.00 (2347)
	CGA	0.27 (475)	1.52 (8672)	**Trp**	UGG	1.00 (2035)	1.00 (5133)
	CGG	0.06 (107)	0.91 (5219)	**TER**	UAA	2.60 (436)	0.26 (430)
	AGA	0.01 (13)	0.68 (3884)		UAG	0.14 (23)	0.24 (410)
	AGG	0.03 (57)	0.65 (3709)		UGA	0.26 (44)	2.50 (4204)

Table 18 | Correspondence Analysis Results for the Opisthokont *C. owczarzaki*

Amino Acids	Codons	Highly Biased Genes	Weakly Biased Genes	Amino Acids	Codons	Highly Biased Genes	Weakly Biased Genes
Phe	UUU	0.56 (2844)	0.99 (1922)	**Ser**	UCU*	1.05 (2087)	0.77 (4023)
	UUC*	1.44 (7257)	1.01 (1964)		UCC*	1.35 (2678)	0.56 (2908)
Leu	UUA	0.26 (819)	0.85 (2090)		UCA	0.79 (1560)	1.17 (6101)
	UUG	1.07 (3396)	1.28 (3152)		UCG*	1.09 (2164)	0.92 (4769)
	CUU*	0.80 (2547)	0.67 (1651)		AGU	0.86 (1706)	1.44 (7501)
	CUC*	1.01 (3234)	0.83 (2049)		AGC	0.86 (1716)	1.14 (5963)
	CUA	0.67 (2142)	0.89 (2205)	**Thr**	ACU*	0.84 (2415)	0.76 (4680)
	CUG*	2.19 (6994)	1.49 (3675)		ACC*	1.63 (4695)	0.74 (4551)
Val	GUU	0.72 (2900)	0.69 (2353)		ACA	0.76 (2201)	1.59 (9722)
	GUC*	1.03 (4145)	0.69 (2339)		ACG	0.78 (2239)	0.91 (5558)
	GUA	0.61 (2465)	1.11 (3766)	**Tyr**	UAU	0.50 (1827)	0.96 (2844)
	GUG*	1.65 (6665)	1.51 (5116)		UAC*	1.50 (5512)	1.04 (3061)
Pro	CCU*	0.91 (1725)	0.82 (4301)	**Ala**	GCU*	1.01 (4289)	0.83 (4560)
	CCC*	1.81 (3413)	0.79 (4121)		GCC*	1.63 (6888)	0.76 (4159)
	CCA	0.56 (1055)	1.46 (7622)		GCA	0.54 (2300)	1.45 (7985)
	CCG	0.72 (1349)	0.92 (4825)		GCG	0.81 (3433)	0.96 (5254)
His	CAU	0.51 (1009)	0.82 (4851)	**Ile**	AUU*	0.98 (4153)	0.82 (2251)
	CAC*	1.49 (2951)	1.18 (6941)		AUC*	1.74 (7348)	0.87 (2392)
Gln	CAA	0.53 (1598)	0.98 (11774)		AUA	0.28 (1164)	1.31 (3609)
	CAG*	1.47 (4404)	1.02 (12356)	**Asp**	GAU	0.93 (5154)	1.08 (7588)
Asn	AAU	0.59 (2371)	1.07 (7560)		GAC*	1.07 (5941)	0.92 (6455)
	AAC*	1.41 (5646)	0.93 (6538)	**Glu**	GAA	0.52 (3117)	0.95 (6258)
Lys	AAA	0.42 (2530)	0.94 (5099)		GAG*	1.48 (8769)	1.05 (6875)
	AAG*	1.58 (9596)	1.06 (5741)	**Gly**	GGU*	1.36 (4751)	1.13 (5588)
Cys	UGU	0.90 (1810)	1.07 (1615)		GGC*	1.16 (4055)	1.03 (5090)
	UGC*	1.10 (2217)	0.93 (1405)		GGA*	1.03 (3591)	0.89 (4423)
Arg	CGU*	1.84 (2589)	0.76 (2603)		GGG	0.44 (1530)	0.95 (4684)
	CGC*	1.84 (2593)	0.91 (3121)	**Met**	AUG	1.00 (6469)	1.00 (6947)
	CGA	0.77 (1091)	0.94 (3202)	**Trp**	UGG	1.00 (2685)	1.00 (1885)
	CGG	0.55 (777)	0.83 (2833)	**TER**	UAA	1.57 (356)	0.72 (349)
	AGA	0.58 (815)	1.33 (4542)		UAG	0.79 (180)	0.79 (384)
	AGG	0.41 (582)	1.24 (4225)		UGA	0.64 (145)	1.49 (722)

Table 19 | Correspondence Analysis Results for the Opisthokont *S. arctica*

Amino Acids	Codons	Highly Biased Genes	Weakly Biased Genes	Amino Acids	Codons	Highly Biased Genes	Weakly Biased Genes
Phe	UUU	0.42 (755)	1.25 (2898)	**Ser**	UCU	0.93 (1778)	1.68 (5314)
	UUC*	1.58 (2868)	0.75 (1734)		UCC*	1.21 (2330)	0.81 (2551)
Leu	UUA	0.19 (258)	0.85 (2854)		UCA	0.44 (842)	1.12 (3540)
	UUG	0.44 (610)	0.88 (2928)		UCG*	0.88 (1680)	0.51 (1612)
	CUU	0.85 (1176)	1.49 (4992)		AGU	0.83 (1591)	0.91 (2872)
	CUC*	2.55 (3542)	0.84 (2814)		AGC*	1.71 (3288)	0.97 (3071)
	CUA	0.37 (510)	1.02 (3406)	**Thr**	ACU	0.77 (2106)	1.21 (3523)
	CUG*	1.62 (2254)	0.91 (3043)		ACC*	1.06 (2899)	0.81 (2369)
Val	GUU	0.64 (1511)	1.27 (2845)		ACA	0.98 (2685)	1.43 (4168)
	GUC*	1.79 (4250)	0.80 (1797)		ACG*	1.20 (3296)	0.55 (1615)
	GUA	0.35 (835)	0.97 (2180)	**Tyr**	UAU	0.58 (1348)	1.22 (2850)
	GUG*	1.22 (2899)	0.96 (2155)		UAC*	1.42 (3330)	0.78 (1826)
Pro	CCU	0.81 (1349)	1.26 (2533)	**Ala**	GCU	0.77 (2516)	1.19 (4141)
	CCC*	1.36 (2273)	0.74 (1487)		GCC*	1.42 (4619)	0.79 (2756)
	CCA	0.68 (1140)	1.43 (2863)		GCA	0.85 (2789)	1.53 (5297)
	CCG*	1.15 (1923)	0.57 (1133)		GCG*	0.96 (3128)	0.49 (1690)
His	CAU	0.45 (436)	1.10 (2542)	**Ile**	AUU	0.55 (850)	1.18 (3854)
	CAC*	1.55 (1515)	0.90 (2076)		AUC*	1.78 (2733)	0.76 (2495)
Gln	CAA	0.50(1000)	0.99 (4599)		AUA	0.66 (1013)	1.06 (3447)
	CAG*	1.50 (2965)	1.01 (4715)	**Asp**	GAU	0.77 (3186)	1.18 (5623)
Asn	AAU	0.68 (2181)	1.14 (4232)		GAC*	1.23 (5076)	0.82 (3890)
	AAC*	1.32 (4246)	0.86(3202)	**Glu**	GAA	0.46 (1519)	0.96 (6079)
Lys	AAA	0.36 (1428)	0.84 (4183)		GAG*	1.54 (5031)	1.04 (6526)
	AAG*	1.64 (6551)	1.16 (5784)	**Gly**	GGU	0.55 (2213)	0.97 (1520)
Cys	UGU	0.61 (4555)	1.05 (1143)		GGC*	1.51 (6114)	0.89 (1396)
	UGC*	1.39 (10315)	0.95 (1026)		GGA	0.83 (3338)	1.33 (2072)
Arg	CGU	0.039(337)	0.94 (1645)		GGG*	1.11 (4499)	0.81 (1258)
	CGC*	1.83 (1589)	0.89 (1558)	**Met**	AUG	1.00 (2045)	1.00 (3858)
	CGA	0.31 (273)	0.93 (1622)	**Trp**	UGG	1.00 (957)	1.00 (750)
	CGG*	1.11 (966)	0.80 (1393)	**TER**	UAA	0.56 (68)	0.99 (122)
	AGA	0.69 (600)	01.48(2583)		UAG	1.14 (140)	1.06 (130)
	AGG*	1.66 (1446)	0.95 (1656)		UGA	1.30 (159)	0.95 (116)

Table 20 | Correspondence Analysis Results for the Excavate *G. lamblia*

Amino Acids	Codons	Highly Biased Genes	Weakly Biased Genes	Amino Acids	Codons	Highly Biased Genes	Weakly Biased Genes
Phe	UUU	0.13 (305)	1.70 (11220)	**Ser**	UCU	0.17 (167)	1.77 (5259)
	UUC*	1.87 (4258)	0.30 (1990)		UCC*	2.08 (1903)	0.27 (788)
Leu	UUA	0.04 (65)	3.23 (13712)		UCA	0.19 (169)	2.01 (6008)
	UUG	0.13 (223)	0.70 (2984)		UCG*	1.30 (1194)	0.30 (844)
	CUU	0.16 (263)	0.99 (4159)		AGU	0.12 (118)	1.26 (3765)
	CUC*	2.52 (2076)	0.14 (559)		AGC*	2.13 (1952)	0.39 (1119)
	CUA	0.16 (259)	0.77 (3204)	**Thr**	ACU	0.16 (184)	1.67 (4284)
	CUG*	3.00 (4839)	0.17 (682)		ACC*	2.18 (2407)	0.29 (722)
Val	GUU	0.16 (251)	1.89 (4871)		ACA	0.15 (173)	1.78 (4690)
	GUC*	1.48 (2391)	0.27 (706)		ACG*	1.52 (1691)	0.26 (647)
	GUA	0.15 (254)	1.49 (3793)	**Tyr**	UAU	0.12 (155)	1.71 (8860)
	GUG*	2.21 (3559)	0.35 (866)		UAC*	1.88 (2393)	0.29 (1523)
Pro	CCU	0.25 (292)	1.55 (2063)	**Ala**	GCU	0.20 (512)	1.62 (2467)
	CCC*	1.95 (2244)	0.16 (220)		GCC*	1.88 (4697)	0.27 (423)
	CCA	0.32 (375)	2.05 (2667)		GCA	0.27 (687)	1.86 (2883)
	CCG*	1.47 (1693)	0.24 (301)		GCG*	1.65 (4101)	0.24 (363)
His	CAU	0.17 (168)	2.68 (2923)	**Ile**	AUU	0.21 (242)	1.58 (13160)
	CAC*	1.83 (1834)	0.32 (520)		AUC*	2.58 (3009)	0.24 (1949)
Glu	GAA	0.15 (372)	1.57 (9552)		AUA	0.21 (247)	1.18 (9866)
	GAG*	1.85 (4551)	0.43 (2609)	**Asp**	GAU	0.19 (400)	1.71 (10261)
Asn	AAU	0.14 (186)	1.68 (14694)		GAC*	1.81 (3821)	0.29 (1694)
	AAC*	1.86 (2442)	0.32 (2803)	**Gln**	CAA	0.14 (194)	0.50 (3546)
Lys	AAA	0.16 (252)	1.58 (13760)		CAG*	3.64 (5203)	0.15 (1079)
	AAG*	1.84 (3009)	0.42 (3591)		UAA	0.03 (45)	2.55 (18027)
Cys	UGU	0.12 (126)	1.49 (2340)		UAG	0.20 (281)	0.80 (5674)
	UGC*	1.88 (1898)	0.51 (776)	**Gly**	GGU	0.23 (337)	1.58 (2126)
Arg	CGU*	0.19 (205)	0.63 (565)		GGC*	2.47 (3516)	0.35 (463)
	CGC	3.03 (3247)	0.18 (146)		GGA	0.22 (330)	1.64 (2220)
	CGA	0.13 (142)	0.79 (674)		GGG*	1,08 (1532)	0.43 (550)
	CGG*	1.74 (1868)	0.16 (137)	**Met**	AUG	1.00 (1907)	1.00 (3777)
	AGA	0.15 (182)	3.65 (3192)	**Trp**	UGG	1.00 (856)	1.00 (918)
	AGG*	0.75 (803)	0.58 (537)	**TER**	UGA	1.00 (394)	1.00 (416)

Table 21 | Correspondence Analysis Results for the Excavate *S. salmonicida*

Amino Acids	Codons	Highly Biased Genes	Weakly Biased Genes	Amino Acids	Codons	Highly Biased Genes	Weakly Biased Genes
Phe	UUU	0.28 (722)	0.90 (2416)	**Ser**	UCU	0.57 (677)	0.89 (7517)
	UUC*	1.72 (4419)	1.10 (2927)		UCC*	1.64 (1954)	1.12 (9441)
Leu	UUA	0.02 (28)	0.27 (990)		UCA	0.11 (133)	0.62 (5256)
	UUG	0.25 (430)	0.94 (3397)		UCG*	1.81 (2159)	1.43 (12079)
	CUU	0.42 (738)	1.11 (4012)		AGU	0.17 (200)	0.38 (3181)
	CUC*	1.51 (2628)	1.40 (5050)		AGC*	1.69 (2017)	1.57 (13308)
	CUA	0.12 (214)	0.56 (2033)	**Thr**	ACU	0.23 (414)	0.68 (3346)
	CUG*	3.68 (6386)	1.70 (6145)		ACC*	1.20 (2142)	1.03 (5041)
Val	GUU	0.31 (768)	0.70 (2917)		ACA	0.25 (449)	0.98 (4794)
	GUC	0.94 (2320)	1.05 (4348)		ACG*	2.32 (4134)	1.31 (6425)
	GUA	0.11 (279)	0.61 (2514)	**Tyr**	UAU	0.10 (224)	0.85 (1594)
	GUG*	2.64 (6540)	1.65 (6842)		UAC*	1.90 (4136)	1.15 (2149)
Pro	CCU	0.32 (411)	0.80 (5583)	**Ala**	GCU	0.44 (1175)	0.78 (8693)
	CCC*	0.97 (1262)	0.68 (4743)		GCC*	1.49 (3969)	0.79 (8755)
	CCA	0.24 (306)	0.87 (6082)		GCA	0.19 (508)	0.91 (10112)
	CCG*	2.48 (3215)	1.64 (11431)		GCG*	1.88 (5019)	1.52 (16906)
His	CAU	0.19 (274)	0.84 (4621)	**Ile**	AUU	0.49 (1095)	0.84 (1267)
	CAC*	1.81 (2575)	1.16 (6369)		AUC*	2.45 (5442)	1.53 (2303)
Gln	CAA	0.08 (213)	0.83 (4743)		AUA	0.06 (129)	0.63 (945)
	CAG*	1.92 (4905)	1.17 (6752)	**Asp**	GAU	0.41 (1488)	0.82 (4397)
Asn	AAU	0.15 (393)	0.64 (1301)		GAC*	1.59 (5799)	1.18 (6306)
	AAC*	1.85 (4721)	1.36 (2738)	**Glu**	GAA	0.13 (659)	0.82 (4755)
Lys	AAA	0.07 (377)	0.70 (1615)		GAG*	1.87 (9461)	1.18 (6880)
	AAG*	1.93 (10083)	1.30 (3011)	**Gly**	GGU	0.77 (1620)	0.85 (4837)
Cys	UGU	0.13 (132)	0.73 (3634)		GGC*	2.83 (5972)	1.47 (8381)
	UGC*	1.87 (1964)	1.27 (6289)		GGA	0.11 (223)	0.95 (5382)
Arg	CGU	0.69 (992)	0.95 (5589)		GGG	0.30 (624)	0.73 (4130)
	CGC*	4.60 (6649)	1.54 (9034)	**Met**	AUG	1.00 (4253)	1.00 (3631)
	CGA	0.09 (134)	1.13 (6606)	**Trp**	UGG	1.00 (1268)	1.00 (4628)
	CGG	0.43 (622)	1.15 (6717)	**TER**	UAA	1.41 (195)	0.35 (437)
	AGA	0.03 (48)	0.54 (3171)		UAG	0.97 (134)	0.41 (509)
	AGG	0.15 (221)	0.69 (4034)		UGA	0.62 (86)	2.24 (2775)

Table 21 | Correspondence Analysis Results for the Excavate *L. major*

Amino Acids	Codons	Highly Biased Genes	Weakly Biased Genes	Amino Acids	Codons	Highly Biased Genes	Weakly Biased Genes
Phe	UUU	0.21(895)	1.01 (1238)	**Ser**	UCU	0.28 (599)	1.04 (3595)
	UUC*	1.79(7790)	0.99 (1211)		UCC*	1.71 (3630)	0.79 (2719)
Leu	UUA	0.02(54)	0.26 (805)		UCA	0.07 (139)	1.17 (4029)
	UUG	0.25(736)	0.45 (1375)		UCG*	1.92 (4083)	1.35 (4648)
	CUU	0.23(693)	1.87 (5701)		AGU	0.15 (325)	0.60 (2062)
	CUC*	1.52(4511)	1.19 (3635)		AGC*	1.87 (3968)	1.05 (3626)
	CUA	0.05(141)	1.24 (3792)	**Thr**	ACU	0.13 (399)	0.81 (2575)
	CUG*	3.94(11715)	0.99 (3008)		ACC*	1.31 (4175)	0.81 (2572)
Val	GUU	0.20(889)	1.04 (3793)		ACA	0.10 (321)	1.11 (3509)
	GUC	0.93(4038)	1.02 (3709)		ACG*	2.46 (7831)	1.26 (3995)
	GUA	0.05(225)	0.89 (3262)	**Tyr**	UAU	0.06(224)	1.28 (1164)
	GUG*	2.82(12273)	1.05 (3823)		UAC*	1.94(7096)	0.72 (658)
Pro	CCU	0.23 (543)	1.07 (6513)	**Ala**	GCU	0.32 (1674)	1.11 (9519)
	CCC*	1.24 (2990)	0.81 (4916)		GCC*	1.61 (8433)	0.70 (5998)
	CCA	0.10 (234)	0.97 (5884)		GCA	0.12 (625)	1.13 (9629)
	CCG*	2.43 (5860)	1.14 (6932)		GCG*	1.95 (10237)	1.06 (9052)
His	CAU	0.09(227)	1.12 (7575)	**Ile**	AUU	0.42(1561)	1.00 (567)
	CAC*	1.91(4855)	0.88 (5895)		AUC*	2.55(9425)	1.32 (754)
Gln	CAA	0.07(283)	1.51 (8724)		AUA	0.02(87)	0.68 (387)
	CAG*	1.93(8073)	0.49 (2838)	**Asp**	GAU	0.30(1838)	1.14 (5861)
Asn	AAU	0.08(387)	1.04 (886)		GAC*	1.70(10499)	0.86 (4452)
	AAC*	1.92(8991)	0.96 (824)	**Glu**	GAA	0.08(608)	1.63 (9630)
Lys	AAA	0.05(421)	1.33 (1517)		GAG*	1.92(14464)	0.37 (2182)
	AAG*	1.95(16428)	0.67 (756)	**Gly**	GGU	0.66 (2587)	0.79 (5952)
Cys	UGU	0.10 (173)	0.77 (3954)		GGC*	2.97 (11571)	1.00 (7532)
	UGC*	1.90 (3234)	1.23 (6315)		GGA	0.08 (328)	1.63 (12317)
Arg	CGU	0.68 (1521)	0.95 (8383)		GGG	0.29 (1117)	0.58 (4406)
	CGC*	4.78 (10678)	1.21 (10680)	**Met**	AUG	1.00(6811)	1.00 (955)
	CGA	0.03 (77)	1.55 (13685)	**Trp**	UGG	1.00 (2400)	1.00 (5065)
	CGG	0.42 (928)	1.13 (9968)	**TER**	UAA	1.88(431)	0.38 (899)
	AGA	0.01 (15)	0.58 (5164)		UAG	0.78(180)	0.23 (534)
	AGG	0.08 (189)	0.59 (5196)		UGA	0.34 (78)	2.39 (5631)

Table 22 | Correspondence Analysis Results for the Excavate *L. pyrrhocoris*

Phe	UUU	0.44 (1351)	0.90 (1986)	**Ser**	UCU	0.52 (765)	1.03 (3855)
	UUC*	1.56 (4847)	1.10 (2428)		UCC*	1.50 (2210)	0.76 (2863)
Leu	UUA	0.06 (131)	0.38 (1180)		UCA	0.22 (321)	1.03 (3859)
	UUG	0.45 (956)	0.85 (2661)		UCG*	1.57 (2316)	1.11 (4169)
	CUU	0.44 (941)	1.53 (4814)		AGU	0.26 (390)	0.69 (2596)
	CUC*	1.69 (3587)	1.26 (3965)		AGC*	1.94 (2862)	1.38 (5198)
	CUA	0.16 (330)	0.86 (2697)	**Thr**	ACU	0.28 (567)	0.78 (2813)
	CUG*	3.19 (6760)	1.13 (3545)		ACC*	1.41 (2877)	0.93 (3371)
Val	GUU	0.34 (990)	0.89 (2666)		ACA	0.32 (653)	1.05 (3801)
	GUC	1.02 (2966)	1.14 (3438)		ACG*	1.99 (4070)	1.24 (4497)
	GUA	0.17 (488)	0.72 (2151)	**Tyr**	UAU	0.18 (449)	0.99 (1439)
	GUG*	2.47 (7157)	1.25 (3756)		UAC*	1.82 (4617)	1.01 (1455)
Pro	CCU	0.38 (620)	1.04 (4907)	**Ala**	GCU	0.51 (1688)	1.06 (7527)
	CCC*	1.18 (1926)	0.84 (3954)		GCC*	1.59 (5295)	0.75 (5344)
	CCA	0.33 (531)	1.00 (4698)		GCA	0.33 (1085)	1.16 (8230)
	CCG*	2.11 (3432)	1.13 (5325)		GCG*	1.58 (5248)	1.03 (7315)
His	CAU	0.21 (352)	0.97 (5025)	**Ile**	AUU	0.77 (1990)	0.87 (1064)
	CAC*	1.79 (2970)	1.03 (5388)		AUC*	2.17 (5584)	1.43 (1752)
Gln	CAA	0.16 (441)	1.16 (5548)		AUA	0.06 (150)	0.71 (867)
	CAG*	1.84 (5118)	0.84 (4046)	**Asp**	GAU	0.52 (2165)	1.03 (3538)
Asn	AAU	0.24 (723)	0.88 (1141)		GAC*	1.48 (6149)	0.97 (3311)
	AAC*	1.76 (5426)	1.12 (1439)	**Glu**	GAA	0.20 (1069)	1.34 (4916)
Lys	AAA	0.11 (594)	1.05 (2186)		GAG*	1.80 (9537)	0.66 (2419)
	AAG*	1.89 (10016)	0.95 (1975)	**Gly**	GGU*	0.92 (2291)	0.79 (4186)
Cys	UGU	0.20 (238)	0.72 (4249)		GGC*	2.62 (6549)	1.15 (6048)
	UGC*	1.80 (2191)	1.28 (7604)		GGA	0.16 (390)	1.36 (7189)
Arg	CGU*	1.04 (1500)	0.86 (5559)		GGG	0.30 (757)	0.70 (3692)
	CGC*	4.16 (5995)	1.26 (8105)	**Met**	AUG	1.00 (4722)	1.00 (2011)
	CGA	0.14 (202)	1.23 (7913)	**Trp**	UGG	1.00 (1688)	1.00 (5920)
	CGG	0.48 (687)	1.10 (7094)	**TER**	UAA	1.74 (232)	0.45 (1164)
	AGA	0.02 (29)	0.75 (4798)		UAG	0.69 (92)	0.44 (1146)
	AGG	0.16 (233)	0.80 (5120)		UGA	0.57 (76)	2.12 (5526)

Table 23 | Correspondence Analysis Results for the Excavate *L. seymouri*

Amino Acids	Codons	Highly Biased Genes	Weakly Biased Genes	Amino Acids	Codons	Highly Biased Genes	Weakly Biased Genes
Phe	UUU	0.84 (1632)	1.21 (4752)	**Ser**	UCU	1.00 (1645)	0.197 (2052)
	UUC*	1.16 (2251)	0.79 (3085)		UCC*	1.01 (1664)	0.92 (1941)
Leu	UUA	0.45 (1170)	108 (2884)		UCA	0.94 (1548)	1.22 (2576)
	UUG	1.09 (2847)	1.94 (5190)		UCG	0.82 (1359)	1.03 (2173)
	CUU*	1.64 (4269)	0.74 (1967)		AGU*	1.08 (1784)	0.83 (1765)
	CUC*	1.00 (2613)	0.65 (1741)		AGC*	1.15 (1898)	1.03 (2189)
	CUA*	0.51 (1339)	0.42 (1112)	**Thr**	ACU*	1.01 (2241)	0.81 (1806)
	CUG*	1.31 (3413)	1.18 (3149)		ACC*	0.97 (2162)	0.91 (2011)
Val	GUU*	1.12 (2930)	0.88 (1996)		ACA	1.00 (2216)	1.25 (2769)
	GUC	0.59 (1535)	0.64 (1459)		ACG	1.02 (2265)	1.03 (2279)
	GUA	0.58 (1510)	0.73 (1651)	**Tyr**	UAU	0.74 (1202)	1.16 (2085)
	GUG	1.71 (4467)	1.75 (3990)		UAC*	1.26 (2029)	0.84 (1506)
Pro	CCU*	1.01 (1406)	0.88 (1475)	**Ala**	GCU*	1.21 (5370)	0.86 (1305)
	CCC	0.88 (1227)	0.87 (1456)		GCC*	0.88 (3917)	0.78 (1187)
	CCA	1.10 (1533)	1.18 (1975)		GCA	1.00 (4429)	1.23 (1860)
	CCG	1.01 (1414)	1.07 (1793)		GCG	0.92 (4075)	1.13 (1705)
His	CAU	0.82 (1429)	0.91 (2071)	**Ile**	AUU*	1.36 (2960)	1.11 (2295)
	CAC*	1.18 (2039)	1.09 (2487)		AUC*	1.12 (2448)	0.84 (1725)
Gln	CAA	0.77 (3932)	1.03 (1896)		AUA	0.52 (1125)	1.05 (2177)
	CAG*	1.23 (6251)	0.97 (1771)	**Asp**	GAU*	1.08 (5790)	1.08 (697)
Asn	AAU	0.93 (2942)	0.93 (1659)		GAC	0.92 (4900)	0.92 (858)
	AAC	1.07 (3395)	1.07 (1913)	**Glu**	GAA	0.85 (9952)	1.16 (1333)
Lys	AAA	0.64 (4189)	1.10 (2593)		GAG*	1.15 (13411)	0.84 (971)
	AAG*	1.36 (8949)	0.90 (2101)	**Gly**	GGU*	1.57 (3169)	0.87 (1317)
Cys	UGU	0.78 (709)	1.02 (3558)		GGC*	0.97 (1942)	0.78 (1175)
	UGC*	1.22 (1107)	0.98 (3424)		GGA	0.79 (1586)	1.16 (1759)
Arg	CGU*	1.62 (3950)	0.66 (1459)		GGG	0.67 (1352)	1.19 (1797)
	CGC*	1.74 (4223)	0.74 (1634)	**Met**	AUG	1.00 (4423)	1.00 (4280)
	CGA	0.66 (1601)	0.76 (1685)	**Trp**	UGG	1.00 (929)	1.00 (3520)
	CGG	0.90 (2195)	0.90 (1989)	**TER**	UAA	1.28 (186)	0.84 (1545)
	AGA	0.39 (959)	1.21 (2666)		UAG	0.93 (135)	0.53 (981)
	AGG	0.69 (1671)	1.73 (3814)		UGA	0.79 (115)	1.63 (2998)

Table 24 | Correspondence Analysis Results for the Excavate *T. brucei*

Amino Acids	Codons	Highly Biased Genes	Weakly Biased Genes	Amino Acids	Codons	Highly Biased Genes	Weakly Biased Genes
Phe	UUU*	1.27 (9066)	0.94 (7858)	**Ser**	UCU	0.75 (3507)	0.94 (8124)
	UUC	0.73 (5237)	1.06 (8796)		UCC*	1.15 (5345)	0.93 (8008)
Leu	UUA	0.24 (1794)	0.60 (3657)		UCA	0.65 (3038)	1.20 (10284)
	UUG	1.07 (8164)	2.06 (12600)		UCG*	1.19 (5549)	0.90 (7757)
	CUU*	1.36 (10351)	0.75 (4565)		AGU*	0.88 (4073)	0.79 (6776)
	CUC*	1.02 (7767)	0.93 (5666)		AGC*	1.38 (6399)	1.24 (10638)
	CUA	0.25 (1909)	0.41 (2480)	**Thr**	ACU	0.53 (3115)	0.80 (8028)
	CUG*	2.07 (15789)	1.26 (7717)		ACC	0.79 (4694)	0.84 (8503)
Val	GUU	0.80 (6893)	0.92 (5728)		ACA	0.98 (5796)	1.11 (11165)
	GUC	0.69 (5972)	0.88 (5441)		ACG*	1.70 (10098)	1.25 (12558)
	GUA	0.30 (2562)	0.58 (3596)	**Tyr**	UAU	0.49 (2762)	1.04 (4349)
	GUG*	2.21 (19066)	1.62 (10022)		UAC*	1.51 (8507)	0.96 (4037)
Pro	CCU	0.63 (2613)	0.82 (6096)	**Ala**	GCU	0.58 (6405)	0.88 (10450)
	CCC*	0.97 (4014)	0.72 (5302)		GCC*	1.12 (12225)	0.67 (7962)
	CCA	0.85 (3525)	1.05 (7760)		GCA	0.78 (8601)	1.17 (13858)
	CCG*	1.54 (6348)	1.41 (10408)		GCG*	1.52 (16602)	1.27 (14965)
His	CAU	0.75 (3569)	0.91 (8306)	**Ile**	AUU*	1.56 (9396)	1.09 (4680)
	CAC*	1.25 (5907)	1.09 (10036)		AUC	1.11 (6697)	1.28 (5499)
Gln	CAA	0.49 (5363)	1.21 (9692)		AUA	0.33 (1980)	0.63 (2713)
	CAG*	1.51 (16310)	0.79 (6373)	**Asp**	GAU	0.87 (11875)	0.95 (5451)
Asn	AAU	0.82 (6995)	0.82 (4514)		GAC*	1.13 (15533)	1.05 (5981)
	AAC	1.18 (10151)	1.18 (6527)	**Glu**	GAA	0.57 (13557)	1.33 (9325)
Lys	AAA	0.51 (7790)	1.38 (10409)		GAG*	1.43 (34038)	0.67 (4726)
	AAG*	1.49 (22881)	0.62 (4705)	**Gly**	GGU*	0.94 (6441)	0.72 (8614)
Cys	UGU	0.67 (1879)	0.84 (16670)		GGC*	1.57 (10789)	0.98 (11782)
	UGC*	1.33 (3702)	1.16 (23092)		GGA	0.59 (4020)	1.58 (19006)
Arg	CGU*	1.43 (8464)	0.72 (10080)		GGG*	0.90 (6195)	0.73 (8783)
	CGC*	1.88 (11146)	0.96 (13347)	**Met**	AUG	1.00 (12497)	1.00 (6729)
	CGA	0.62 (3658)	1.08 (15086)	**Trp**	UGG	1.00 (4548)	1.00 (19899)
	CGG*	1.19 (7088)	1.09 (15179)	**TER**	UAA	0.88 (355)	0.53 (4894)
	AGA	0.25 (1510)	1.14 (15830)		UAG	0.80 (325)	0.23 (2160)
	AGG	0.63 (3762)	1.00 (13907)		UGA	1.32 (532)	2.24 (20646)

Table 25 | Correspondence Analysis Results for the Excavate *T. cruzi*

Amino Acids	Codons	Highly Biased Genes	Weakly Biased Genes	Amino Acids	Codons	Highly Biased Genes	Weakly Biased Genes
Phe	UUU*	1.70 (18453)	0.74 (1758)	**Ser**	UCU*	2.10 (17565)	0.76 (1101)
	UUC	0.30 (3214)	1.26 (2995)		UCC	0.51 (4293)	1.27 (1846)
Leu	UUA*	2.35 (22563)	0.24 (465)		UCA*	1.37 (11415)	0.68 (991)
	UUG	0.70 (6708)	0.83 (1606)		UCG	0.18 (1508)	1.05 (1524)
	CUU*	1.89 (18136)	0.80 (1562)		AGU*	1.69 (14133)	0.81 (1182)
	CUC	0.43 (4102)	1.87 (3627)		AGC	0.14 (1179)	1.43 (2086)
	CUA*	0.41 (3954)	0.23 (438)	**Thr**	ACU*	1.39 (12371)	0.52 (997)
	CUG	0.23 (2181)	2.04 (3953)		ACC	0.43 (3865)	1.13 (2159)
Val	GUU*	1.53 (14282)	0.64 (1645)		ACA*	1.78 (15903)	1.11 (2116)
	GUC	0.23 (2147)	0.83 (2145)		ACG	0.40 (3588)	1.24 (2382)
	GUA*	1.46 (13613)	0.30 (767)	**Tyr**	UAU*	1.63 (12724)	0.43 (765)
	GUG	0.78 (7267)	2.24 (5784)		UAC	0.37 (2922)	1.57 (2754)
Pro	CCU*	1.60 (9740)	0.52 (909)	**Ala**	GCU*	1.47 (12283)	0.57 (1852)
	CCC	0.35 (2124)	1.14 (2001)		GCC	0.61 (5126)	1.44 (4653)
	CCA*	1.79 (10932)	1.01 (1773)		GCA*	1.47 (12229)	0.75 (2418)
	CCG	0.26 (1613)	1.32 (2316)		GCG	0.45 (3746)	1.24 (4024)
His	CAU*	1.68 (12064)	0.43 (749)	**Ile**	AUU*	1.83 (19342)	1.18 (2083)
	CAC	0.32 (2305)	1.57 (2749)		AUC	0.25 (2649)	1.52 (2681)
Gln	CAA*	1.56 (17410)	0.48 (1217)		AUA*	0.92 (9676)	0.29 (513)
	CAG	0.44 (4868)	1.52 (3811)	**Asp**	GAU*	1.84 (25157)	0.75 (2347)
Asn	AAU*	1.77 (28461)	0.58 (1447)		GAC	0.16 (2155)	1.25 (3902)
	AAC	0.23 (3633)	1.42 (3556)	**Glu**	GAA*	1.51 (31011)	0.47 (1944)
Lys	AAA*	1.46 (20370)	0.31 (1226)		GAG	0.49 (9995)	1.53 (6354)
	AAG	0.54 (7573)	1.69 (6744)	**Gly**	GGU*	1.67 (12175)	1.09 (2567)
Cys	UGU*	1.80 (8594)	0.61 (807)		GGC	0.15 (1064)	1.66 (3917)
	UGC	0.20 (979)	1.39 (1828)		GGA*	1.81 (13242)	0.60 (1406)
Arg	CGU*	2.38 (13395)	1.10 (2076)		GGG	0.38 (2762)	0.65 (1539)
	CGC	0.26 (1453)	2.54 (4810)	**Met**	AUG	1.00 (13741)	1.00 (3682)
	CGA*	1.54 (8655)	0.45 (857)	**Trp**	UGG	1.00 (6365)	1.00 (1785)
	CGG	0.35 (1975)	1.07 (2027)	**TER**	UAA	1.42 (246)	1.11 (311)
	AGA*	1.20 (6749)	0.32 (602)		UAG	0.73 (127)	0.85 (237)
	AGG	0.26 (1487)	0.52 (980)		UGA	0.85 (148)	1.04 (292)

Table 26 | Correspondence Analysis Results for the Excavate *T. theileri*

Amino Acids	Codons	Highly Biased Genes	Weakly Biased Genes	Amino Acids	Codons	Highly Biased Genes	Weakly Biased Genes
Phe	UUU	1.18 (14570)	1.46 (28208)	**Ser**	UCU*	2.21 (20766)	1.10 (9813)
	UUC*	0.82 (10062)	0.54 (10387)		UCC	0.35 (3260)	0.53 (4783)
Leu	UUA*	2.59 (19002)	1.48 (30770)		UCA*	1.56 (14643)	1.44 (12903)
	UUG	0.91 (6708)	1.20 (24956)		UCG	0.08 (746)	0.47 (4232)
	CUU*	1.45 (10623)	0.95 (19803)		AGU	1.17 (11010)	1.34 (11997)
	CUC*	0.57 (4177)	0.51 (10651)		AGC	0.62 (5852)	1.11 (9962)
	CUA	0.41 (3027)	1.03 (21396)	**Thr**	ACU*	2.30 (23490)	1.37 (13555)
	CUG	0.08 (557)	0.81 (16896)		ACC	0.46 (4662)	0.47 (4617)
Val	GUU*	2.28 (16093)	1.03 (8962)		ACA	1.20 (12239)	1.62 (16006)
	GUC*	0.62 (4342)	0.44 (3831)		ACG	0.05 (460)	0.55 (5461)
	GUA	0.95 (6682)	1.31 (11379)	**Tyr**	UAU	1.36 (17471)	1.40 (16193)
	GUG	0.16 (1120)	1.21 (10484)		UAC*	0.64 (8172)	0.60 (6918)
Pro	CCU*	2.44 (13994)	1.44 (4677)	**Ala**	GCU*	2.47 (20611)	1.47 (6866)
	CCC	0.43 (2435)	0.59 (1916)		GCC*	0.45 (3752)	0.39 (1831)
	CCA	1.09 (6252)	1.59 (5179)		GCA	1.04 (8704)	1.64 (7637)
	CCG	0.04 (233)	0.38 (1232)		GCG	0.03 (253)	0.50 (2341)
His	CAU	1.21 (3887)	1.45 (11449)	**Ile**	AUU*	1.77 (20387)	1.14 (32905)
	CAC*	0.79 (2518)	0.55 (4372)		AUC*	0.53 (6072)	0.47 (13562)
Glu	GAA*	1.79 (17850)	1.16 (12916)		AUA	0.70 (8117)	1.39 (40017)
	GAG	0.21 (2093)	0.84 (9443)	**Asp**	GAU*	1.62 (24531)	1.48 (8099)
Asn	AAU	1.46 (36698)	1.51 (37359)		GAC	0.38 (5768)	0.52 (2820)
	AAC*	0.54 (13520)	0.49 (12263)	**Gln**	CAA*	0.84 (8293)	1.01 (20269)
Lys	AAA*	1.32 (22566)	1.27 (64793)		CAG	0.10 (957)	0.61 (12312)
	AAG	0.68 (11623)	0.73 (37607)		UAA*	2.57 (25488)	1.42 (28690)
Cys	UGU*	1.18 (24493)	1.12 (6853)		UAG	0.50 (4957)	0.96 (19357)
	UGC	0.82 (16888)	0.88 (5372)	**Gly**	GGU*	2.10 (21405)	1.00 (2386)
Arg	CGU*	0.82 (1586)	0.19 (1000)		GGC	0.38 (3895)	0.52 (1240)
	CGC	0.09 (176)	0.14 (743)		GGA	1.39 (14105)	1.53 (3655)
	CGA	0.07 (128)	0.41 (2199)		GGG	0.13 (1314)	0.95 (2275)
	CGG	0.01 (21)	0.13 (687)	**Met**	AUG	1.00 (8383)	1.00 (28831)
	AGA*	4.67 (9014)	3.37 (18174)	**Trp**	UGG	1.00 (4315)	1.00 (7615)
	AGG	0.34 (656)	1.77 (9567)	**TER**	UGA	1.00 (1231)	1.00 (15631)

Table 27 | Correspondence Analysis Results for the Alveolate *T. thermophila*

Amino Acids	Codons	Highly Biased Genes	Weakly Biased Genes	Amino Acids	Codons	Highly Biased Genes	Weakly Biased Genes
Phe	UUU*	1.56 (35202)	1.42 (11873)	Ser	UCU*	1.52 (18436)	1.40 (11495)
	UUC	0.44 (10058)	0.58 (4819)		UCC	0.23 (2753)	0.47 (3843)
Leu	UUA*	2.48 (27837)	1.89 (20577)		UCA*	2.05 (24928)	1.71 (14063)
	UUG*	0.86 (9666)	0.83 (9018)		UCG	0.20 (2383)	0.28 (2299)
	CUU	1.65 (18559)	1.72 (18821)		AGU*	1.49 (18115)	1.38 (11404)
	CUC	0.26 (2898)	0.49 (5321)		AGC	0.52 (6257)	0.77 (6302)
	CUA	0.60 (6692)	0.80 (8770)	Thr	ACU*	1.82 (23284)	1.70 (11140)
	CUG	0.16 (1793)	0.27 (2981)		ACC	0.31 (4014)	0.60 (3939)
Val	GUU*	2.02 (21222)	1.90 (10370)		ACA*	1.74 (22254)	1.52 (9972)
	GUC	0.41 (4308)	0.52 (2812)		ACG	0.13 (1647)	0.19 (1236)
	GUA*	1.11 (11632)	1.01 (5476)	Tyr	UAU*	1.63 (36090)	1.50 (12128)
	GUG	0.46 (4842)	0.57 (3129)		UAC	0.37 (8201)	0.50 (4049)
Pro	CCU*	1.98 (13712)	1.79 (6073)	Ala	GCU	2.11 (19465)	2.08 (13966)
	CCC	0.12 (862)	0.26 (879)		GCC	0.29 (2714)	0.40 (2698)
	CCA	1.74 (12083)	1.70 (5764)		GCA*	1.40 (12914)	1.19 (7962)
	CCG	0.16 (1090)	0.25 (835)		GCG	0.20 (1838)	0.33 (2189)
His	CAU*	1.73 (7845)	1.61 (8114)	Ile	AUU*	1.44 (32912)	1.16 (18064)
	CAC	0.27 (1246)	0.39 (1947)		AUC	0.41 (9342)	0.66 (10341)
Gln	CAA*	1.60 (15559)	1.51 (29064)		AUA	1.15 (26435)	1.18 (18370)
	CAG	0.40 (3873)	0.49 (9478)	Asp	GAU*	1.65 (27798)	1.49 (24984)
Asn	AAU*	1.53 (34619)	1.34 (29366)		GAC	0.35 (5828)	0.51 (8636)
	AAC	0.47 (10768)	0.66 (14383)	Glu	GAA*	1.64 (25454)	1.61 (75557)
Lys	AAA*	1.65 (28859)	1.63 (78926)		GAG	0.36 (5676)	0.39 (18266)
	AAG	0.35 (6176)	0.37 (17932)	Gly	GGU*	1.18 (15007)	1.04 (3317)
Cys	UGU*	1.36 (22912)	1.26 (2759)		GGC	0.36 (4581)	0.43 (1364)
	UGC	0.64 (10731)	0.74 (1615)		GGA*	1.67 (21234)	1.39 (4425)
Arg	CGU	0.76 (1931)	0.83 (4505)		GGG	0.79 (10015)	1.14 (3612)
	CGC	0.15 (386)	0.23 (1281)	Met	AUG	1.00 (16247)	1.00 (15119)
	CGA	0.48 (1214)	0.57 (3115)	Trp	UGG	1.00 (9898)	1.00 (2304)
	CGG	0.10 (245)	0.18 (988)	TER	UAA	1.73 (908)	2.16 (1131)
	AGA*	3.44 (8720)	2.82 (15388)		UAG	1.27 (663)	0.84 (440)
	AGG	1.08 (2734)	1.37 (7462)		UGA	0.00 (0)	0.00 (0)

Table 28 | Correspondence Analysis Results for the Alveolate *S. coeruleus*

Amino Acids	Codons	Highly Biased Genes	Weakly Biased Genes	Amino Acids	Codons	Highly Biased Genes	Weakly Biased Genes
Phe	UUU*	1.79 (6725)	1.43 (6020)	**Ser**	UCU*	1.40 (3989)	0.75 (527)
	UUC	0.21 (777)	0.57 (2429)		UCC	0.23 (664)	0.41 (285)
Leu	UUA*	4.21 (9643)	2.13 (9318)		UCA	1.59 (4517)	1.75 (1220)
	UUG	0.68 (1569)	1.15 (5032)		UCG	0.14 (396)	0.50 (346)
	CUU	0.54 (1231)	0.68 (2988)		AGU*	2.24 (6387)	1.69 (1181)
	CUC	0.06 (127)	0.33 (1434)		AGC	0.40 (1145)	0.90 (632)
	CUA	0.44 (1003)	1.14 (5003)	**Thr**	ACU*	1.58 (3493)	0.75 (1667)
	CUG	0.08 (173)	0.56 (2462)		ACC	0.24 (530)	0.25 (569)
Val	GUU*	1.79 (3101)	0.86 (2350)		ACA	1.99 (4384)	2.36 (5264)
	GUC	0.15 (254)	0.24 (647)		ACG	0.19 (427)	0.64 (1436)
	GUA	1.74 (3004)	1.81 (4978)	**Tyr**	UAU*	1.74 (8398)	1.45 (5633)
	GUG	0.32 (561)	1.09 (3004)		UAC	0.26 (1234)	0.55 (2147)
Pro	CCU*	1.58 (1643)	1.19 (522)	**Ala**	GCU*	1.56 (1607)	0.86 (430)
	CCC	0.23 (244)	0.43 (190)		GCC	0.22 (229)	0.31 (154)
	CCA@	2.06 (2140)	1.93 (848)		GCA	2.11 (2177)	2.33 (1165)
	CCG	0.13 (131)	0.46 (201)		GCG	0.11 (110)	0.51 (255)
His	CAU*	1.75 (3098)	1.44 (3366)	**Ile**	AUU*	1.39 (8070)	0.77 (12856)
	CAC	0.25 (436)	0.56 (1300)		AUC	0.15 (869)	0.23 (3754)
Gln	CAA*	1.74 (4452)	1.33 (6024)		AUA	1.46 (8441)	2.00 (33160)
	CAG	0.26 (660)	0.67 (3035)	**Asp**	GAU*	1.80 (13950)	1.51 (1166)
Asn	AAU*	1.76 (37758)	1.46 (7848)		GAC	0.20 (1512)	0.49 (379)
	AAC	0.24 (5183)	0.54 (2901)	**Glu**	GAA*	1.79 (18230)	1.36 (3670)
Lys	AAA*	1.74 (22047)	1.38 (27271)		GAG	0.21 (2174)	0.64 (1737)
	AAG	0.26 (3230)	0.62 (12370)	**Gly**	GGU*	1.46 (2131)	0.67 (291)
Cys	UGU*	1.61 (2364)	1.23 (1272)		GGC	0.17 (243)	0.37 (160)
	UGC	0.39 (577)	0.77 (788)		GGA	2.10 (3054)	2.06 (900)
Arg	CGU*	0.39 (348)	0.22 (242)		GGG	0.27 (392)	0.91 (397)
	CGC	0.05 (46)	0.09 (98)	**Met**	AUG	1.00 (3851)	1.00 (15827)
	CGA	0.21 (192)	0.53 (582)	**Trp**	UGG	1.00 (749)	1.00 (1341)
	CGG	0.03 (25)	0.20 (220)	**TER**	UAA	2.50 (214)	1.82 (13759)
	AGA*	4.68 (4177)	3.04 (3365)		UAG	0.23 (20)	0.70 (5284)
	AGG	0.64 (572)	1.93 (2137)		UGA	0.27 (23)	0.48 (3591)

Table 29 | Correspondence Analysis Results for the Alveolate *P. gallinaceum*

Amino Acids	Codons	Highly Biased Genes	Weakly Biased Genes	Amino Acids	Codons	Highly Biased Genes	Weakly Biased Genes
Phe	UUU*	1.70 (11650)	1.61 (2014)	**Ser**	UCU	1.24 (5400)	1.41 (1111)
	UUC	0.30 (2095)	0.39 (494)		UCC	0.41 (1773)	0.59 (463)
Leu	UUA	3.66 (14265)	3.56 (4239)		UCA	1.40 (6103)	1.93 (1519)
	UUG	0.86 (3368)	1.02 (1219)		UCG	0.27 (1193)	0.32 (249)
	CUU	0.73 (2830)	0.67 (798)		AGU*	2.27 (9873)	1.36 (1066)
	CUC	0.14 (546)	0.14 (164)		AGC	0.41 (1786)	0.39 (303)
	CUA	0.49 (1902)	0.49 (586)	**Thr**	ACU	0.98 (3734)	1.34 (1756)
	CUG	0.13 (490)	0.11 (135)		ACC	0.40 (1544)	0.50 (651)
Val	GUU	1.45 (4569)	1.99 (3649)		ACA*	2.25 (8594)	1.88 (2473)
	GUC	0.23 (712)	0.28 (504)		ACG*	0.37 (1412)	0.28 (372)
	GUA*	1.69 (5323)	1.30 (2378)	**Tyr**	UAU*	1.80 (21662)	1.57 (2189)
	GUG*	0.63 (1968)	0.43 (786)		UAC	0.20 (2413)	0.43 (592)
Pro	CCU*	1.66 (2284)	1.44 (1277)	**Ala**	GCU	1.55 (1603)	1.91 (3254)
	CCC*	0.41 (572)	0.28 (249)		GCC	0.35 (360)	0.47 (805)
	CCA	1.75 (2415)	2.17 (1922)		GCA*	1.85 (1912)	1.40 (2382)
	CCG*	0.18 (244)	0.11 (96)		GCG	0.25 (259)	0.22 (368)
His	CAU*	1.74 (9995)	1.43 (1096)	**Ile**	AUU	1.15 (13745)	1.43 (3489)
	CAC	0.26 (1510)	0.57 (437)		AUC	0.19 (2320)	0.25 (613)
Gln	CAA	1.74 (9381)	1.78 (2680)		AUA*	1.66 (19884)	1.32 (3235)
	CAG	0.26 (1395)	0.22 (339)	**Asp**	GAU*	1.75 (25339)	1.56 (2798)
Asn	AAU*	1.72 (81443)	1.54 (2853)		GAC	0.25 (3559)	0.44 (787)
	AAC	0.28 (13125)	0.46 (858)	**Glu**	GAA*	1.70 (21807)	1.67 (7519)
Lys	AAA	1.62 (37891)	1.63 (6623)		GAG	0.30 (3789)	0.33 (1462)
	AAG	0.38 (8974)	0.37 (1507)	**Gly**	GGU	1.68 (3338)	2.01 (2438)
Cys	UGU*	1.79 (5329)	1.42 (1260)		GGC*	0.17 (343)	0.13 (155)
	UGC	0.21 (627)	0.58 (516)		GGA*	1.71 (3398)	1.60 (1936)
Arg	CGU	0.62 (841)	1.28 (684)		GGG*	0.43 (851)	0.27 (326)
	CGC	0.08 (114)	0.19 (104)	**Met**	AUG	1.00 (10834)	1.00 (2037)
	CGA	0.47 (640)	0.86 (462)	**Trp**	UGG	1.00 (813)	1.00 (436)
	CGG	0.06 (75)	0.04 (23)	**TER**	UAA	2.18 (194)	1.97 (175)
	AGA*	3.58 (4872)	2.89 (1548)		UAG	0.25 (22)	0.84 (75)
	AGG*	1.20 (1628)	0.73 (388)		UGA	0.57 (51)	0.19 (17)

Table 30 | Correspondence Analysis Results for the Alveolate *P. falciparum*

Amino Acids	Codons	Highly Biased Genes	Weakly Biased Genes	Amino Acids	Codons	Highly Biased Genes	Weakly Biased Genes
Phe	UUU*	1.58 (3225)	0.72 (2463)	**Ser**	UCU*	1.56 (2071)	0.36 (658)
	UUC	0.42 (846)	1.28 (4370)		UCC	0.43 (567)	1.44 (2651)
Leu	UUA*	2.96 (3650)	0.27 (604)		UCA*	1.75 (2315)	0.26 (485)
	UUG	0.86 (1064)	0.94 (2113)		UCG	0.28 (376)	0.85 (1559)
	CUU*	1.09 (1342)	0.38 (860)		AGU*	1.57 (2073)	0.69 (1266)
	CUC	0.21 (257)	1.68 (3784)		AGC	0.41 (544)	2.40 (4403)
	CUA*	0.68 (837)	0.46 (1035)	**Thr**	ACU*	1.51 (1940)	0.37 (534)
	CUG	0.21 (260)	2.27 (5125)		ACC	0.32 (416)	1.61 (2357)
Val	GUU*	1.64 (1717)	0.56 (1175)		ACA*	1.82 (2339)	0.41 (601)
	GUC	0.26 (271)	1.03 (2161)		ACG	0.34 (433)	1.61 (2350)
	GUA*	1.70 (1787)	0.35 (726)	**Tyr**	UAU*	1.59 (4561)	0.39 (1176)
	GUG	0.40 (418)	2.06 (4317)		UAC	0.41 (1182)	1.61 (4796)
Pro	CCU*	1.61 (1490)	0.36 (445)	**Ala**	GCU*	1.43 (1279)	0.46 (955)
	CCC	0.48 (443)	2.07 (2553)		GCC	0.30 (267)	1.56 (3247)
	CCA*	1.66 (1538)	0.58 (719)		GCA*	1.94 (1733)	0.59 (1224)
	CCG	0.26 (242)	0.98 (1206)		GCG	0.33 (294)	1.40 (2906)
His	CAU*	1.64 (1713)	0.41 (779)	**Ile**	AUU*	1.38 (2779)	0.89 (2294)
	CAC	0.36 (381)	1.59 (2997)		AUC	0.25 (510)	1.48 (3798)
Gln	CAA*	1.60 (2346)	0.54 (1242)		AUA*	1.37 (2750)	0.62 (1603)
	CAG	0.40 (580)	1.46 (3391)	**Asp**	GAU*	1.67 (4840)	0.57 (2051)
Asn	AAU*	1.65 (6033)	0.55 (2270)		GAC	0.33 (950)	1.43 (5137)
	AAC	0.35 (1284)	1.45 (5987)	**Glu**	GAA*	1.71 (6002)	0.62 (3102)
Lys	AAA*	1.59 (7364)	0.59 (3090)		GAG	0.29 (999)	1.38 (6910)
	AAG	0.41 (1889)	1.41 (7384)	**Gly**	GGU*	1.48 (1677)	0.41 (1000)
Cys	UGU*	1.48 (1643)	0.36 (516)		GGC	0.29 (333)	1.28 (3154)
	UGC	0.52 (570)	1.64 (2373)		GGA*	1.84 (2087)	0.74 (1824)
Arg	CGU*	1.05 (592)	0.21 (285)		GGG	0.39 (444)	1.56 (3840)
	CGC	0.26 (145)	1.12 (1531)	**Met**	AUG	1.00 (1537)	1.00 (2846)
	CGA*	0.67 (374)	0.53 (719)	**Trp**	UGG	1.00 (698)	1.00 (842)
	CGG	0.13 (72)	1.15 (1570)	**TER**	UAA	2.03 (168)	0.40 (36)
	AGA*	2.93 (1649)	0.82 (1118)		UAG	0.52 (43)	1.16 (104)
	AGG	0.96 (541)	2.19 (2993)		UGA	0.45 (37)	1.44 (129)

Table 31 | Correspondence Analysis Results for the Alveolate *P. vivax*

Amino Acids	Codons	Highly Biased Genes	Weakly Biased Genes	Amino Acids	Codons	Highly Biased Genes	Weakly Biased Genes
Phe	UUU	0.42 (2814)	1.21 (2404)	**Ser**	UCU	0.59 (2193)	1.64 (5208)
	UUC*	1.58 (10654)	0.79 (1575)		UCC*	0.97 (3574)	0.78 (2470)
Leu	UUA	0.13 (608)	0.25 (885)		UCA	0.46 (1690)	0.89 (2842)
	UUG*	0.92 (4392)	0.24 (836)		UCG*	1.97 (7274)	0.94 (2994)
	CUU	0.60 (2886)	2.08 (7318)		AGU	0.71 (2615)	0.76 (2421)
	CUC	1.34 (6399)	1.47 (5183)		AGC*	1.30 (4809)	0.99 (3156)
	CUA	0.37 (1763)	1.11 (3897)	**Thr**	ACU	0.57 (2701)	1.40 (3124)
	CUG*	2.65 (12672)	0.85 (2971)		ACC*	1.03 (4890)	0.87 (1936)
Val	GUU	0.40 (2420)	1.40 (6472)		ACA	0.48 (2280)	0.94 (2113)
	GUC	1.07 (6451)	1.16 (5369)		ACG*	1.91 (9044)	0.79 (1774)
	GUA	0.26 (1597)	0.82 (3816)	**Tyr**	UAU	0.30 (1408)	1.32 (4148)
	GUG*	2.27 (13753)	0.62 (2846)		UAC*	1.70 (8133)	0.68 (2151)
Pro	CCU	0.81 (2659)	1.01 (4998)	**Ala**	GCU	1.05 (7268)	1.45 (10423)
	CCC*	1.08 (3555)	0.89 (4370)		GCC*	1.47 (10204)	0.67 (4796)
	CCA	0.51 (1675)	1.24 (6120)		GCA	0.45 (3098)	1.08 (7784)
	CCG*	1.60 (5272)	0.86 (4254)		GCG*	1.04 (7182)	0.80 (5719)
His	CAU	0.34 (1066)	1.18 (10958)	**Ile**	AUU	0.83 (4243)	1.33 (1540)
	CAC*	1.66 (5194)	0.82 (7594)		AUC*	2.09 (10637)	0.98 (1128)
Gln	CAA	0.38 (2315)	1.68 (17131)		AUA	0.07 (379)	0.69 (800)
	CAG*	1.62 (9780)	0.32 (3317)	**Asp**	GAU	0.48 (4351)	1.06 (7364)
Asn	AAU	0.36 (2172)	1.14 (1490)		GAC*	1.52 (13951)	0.94 (6479)
	AAC*	1.64 (9923)	0.86 (1129)	**Glu**	GAA	0.46 (4891)	1.58 (11099)
Lys	AAA	0.29 (3081)	1.51 (2051)		GAG*	1.54 (16525)	0.42 (2955)
	AAG*	1.71 (18155)	0.49 (671)	**Gly**	GGU*	1.18 (5935)	0.70 (6722)
Cys	UGU	0.64 (1421)	0.96 (6003)		GGC*	2.03 (10192)	1.03 (9850)
	UGC*	1.36 (3015)	1.04 (6534)		GGA	0.55 (2775)	1.60 (15308)
Arg	CGU*	2.11 (5500)	1.11 (8895)		GGG	0.24 (1180)	0.66 (6323)
	CGC*	2.46 (6419)	1.10 (8851)	**Met**	AUG	1.00 (8892)	1.00 (713)
	CGA	0.56 (1466)	1.89 (15189)	**Trp**	UGG	1.00 (3682)	1.00 (5719)
	CGG	0.40 (1053)	1.02 (8169)	**TER**	UAA	0.63 (24)	0.90 (3507)
	AGA	0.23 (607)	0.58 (4642)		UAG	0.44 (17)	0.22 (838)
	AGG	0.24 (627)	0.30 (2427)		UGA	1.93 (74)	1.88 (7334)

Table 32 | Correspondence Analysis Results for the Stramenopile *P. infestans*

Amino Acids	Codons	Highly Biased Genes	Weakly Biased Genes	Amino Acids	Codons	Highly Biased Genes	Weakly Biased Genes
Phe	UUU	0.06 (502)	1.09 (4670)	**Ser**	UCU	0.12 (514)	0.99 (4397)
	UUC*	1.94 (16382)	0.91 (3887)		UCC*	1.29 (5663)	0.71 (3142)
Leu	UUA	0.01 (52)	0.56 (2520)		UCA	0.04 (198)	1.15 (5122)
	UUG	0.12 (665)	1.03 (4647)		UCG*	2.53 (11154)	1.17 (5179)
	CUU	0.12 (675)	1.30 (5874)		AGU	0.13 (560)	0.93 (4144)
	CUC*	2.19 (12306)	1.03 (4678)		AGC*	1.89 (8343)	1.05 (4650)
	CUA	0.04 (199)	0.96 (4359)	**Thr**	ACU	0.11 (607)	1.03 (4796)
	CUG*	3.53 (19846)	1.12 (5053)		ACC*	1.32 (7382)	0.82 (3821)
Val	GUU	0.08 (644)	1.14 (5286)		ACA	0.06 (309)	1.12 (5229)
	GUC*	1.27 (9734)	0.93 (4274)		ACG*	2.52 (14125)	1.03 (4809)
	GUA	0.03 (227)	0.86 (3959)	**Tyr**	UAU	0.04 (232)	1.03 (3669)
	GUG*	2.61 (19946)	1.07 (4952)		UAC*	1.96 (11922)	0.97 (3440)
Pro	CCU	0.17 (738)	1.04 (4458)	**Ala**	GCU	0.28 (2759)	1.32 (8863)
	CCC*	1.87 (8340)	0.67 (2862)		GCC*	2.46 (24477)	0.65 (4347)
	CCA	0.05 (207)	1.39 (5932)		GCA	0.06 (594)	1.11 (7463)
	CCG*	1.92 (8580)	0.90 (3820)		GCG*	1.21 (12007)	0.92 (6157)
His	CAU	0.04 (155)	1.09 (5447)	**Ile**	AUU	0.21 (1275)	1.26 (4629)
	CAC*	1.96 (8463)	0.91 (4574)		AUC*	2.78 (17227)	1.09 (3997)
Gln	CAA	0.06 (451)	1.24 (7924)		AUA	0.01 (55)	0.64 (2360)
	CAG*	1.94 (14024)	0.76 (4861)	**Asp**	GAU	0.06 (657)	1.15 (7715)
Asn	AAU	0.04 (306)	1.08 (4596)		GAC*	1.94 (21054)	0.85 (5675)
	AAC*	1.96 (14313)	0.92 (3912)	**Glu**	GAA	0.06 (724)	1.14 (7562)
Lys	AAA	0.03 (372)	0.96 (5429)		GAG*	1.94 (23896)	0.86 (5708)
	AAG*	1.97 (23754)	1.04 (5920)	**Gly**	GGU	0.29 (2090)	1.03 (5415)
Cys	UGU	0.08 (238)	0.95 (3913)		GGC*	3.43 (24842)	0.90 (4738)
	UGC*	1.92 (5528)	1.05 (4307)		GGA	0.15 (1103)	1.29 (6770)
Arg	CGU	0.57 (1794)	1.06 (5840)		GGG	0.13 (942)	0.78 (4084)
	CGC*	4.78 (15157)	1.10 (6052)	**Met**	AUG	1.00 (10486)	1.00 (5140)
	CGA	0.11 (348)	1.41 (7774)	**Trp**	UGG	1.00 (4672)	1.00 (6240)
	CGG	0.38 (1201)	1.00 (5501)	**TER**	UAA	1.67 (691)	0.51 (666)
	AGA	0.04 (116)	0.81 (4454)		UAG	0.97 (402)	0.39 (505)
	AGG	0.13 (424)	0.61 (3353)		UGA	0.36 (147)	2.11 (2762)

Table 33 | Correspondence Analysis Results for the Stramenopile *P. sojae*

Amino Acids	Codons	Highly Biased Genes	Weakly Biased Genes	Amino Acids	Codons	Highly Biased Genes	Weakly Biased Genes
Phe	UUU*	1.48 (6024)	0.97 (6522)	**Ser**	UCU*	1.07 (3107)	0.71 (4243)
	UUC	0.52 (2114)	1.03 (6879)		UCC	0.51 (1483)	0.63 (3764)
Leu	UUA	0.94 (3505)	0.92 (5079)		UCA	1.21 (3512)	1.20 (7169)
	UUG	1.31 (4886)	1.71 (9416)		UCG*	1.21 (3504)	1.07 (6440)
	CUU*	1.59 (5951)	0.77 (4245)		AGU*	1.21 (3515)	0.85 (5126)
	CUC	0.66 (2455)	0.77 (4256)		AGC	0.79 (2299)	1.54 (9251)
	CUA	0.61 (2288)	0.68 (3743)	**Thr**	ACU*	1.10 (3906)	0.85 (4812)
	CUG	0.89 (3334)	1.16 (6388)		ACC	0.53 (1893)	0.76 (4328)
Val	GUU*	1.14 (5186)	0.78 (3340)		ACA	1.24 (4414)	1.20 (6823)
	GUC	0.75 (3407)	1.01 (4343)		ACG	1.13 (4012)	1.20 (6808)
	GUA	0.79 (3595)	0.82 (3508)	**Tyr**	UAU*	1.03 (2964)	0.98 (4462)
	GUG	1.32 (6025)	1.39 (5954)		UAC	0.97 (2782)	1.02 (4675)
Pro	CCU*	1.24 (3045)	0.63 (1705)	**Ala**	GCU*	1.59 (8806)	0.96 (4287)
	CCC	0.53 (1289)	0.63 (1706)		GCC*	0.70 (3897)	0.62 (2773)
	CCA	1.44 (3518)	1.56 (4231)		GCA	1.06 (5891)	1.31 (5855)
	CCG	0.79 (1932)	1.18 (3184)		GCG	0.65 (3622)	1.11 (4966)
His	CAU*	1.14 (2929)	0.98 (4825)	**Ile**	AUU*	1.79 (7389)	1.08 (5760)
	CAC	0.86 (2232)	1.02 (5039)		AUC	0.82 (3396)	1.10 (5838)
Gln	CAA	1.08 (6219)	1.15 (6936)		AUA	0.38 (1568)	0.82 (4390)
	CAG*	0.92 (5272)	0.85 (5147)	**Asp**	GAU*	1.17 (10602)	0.99 (3219)
Asn	AAU*	1.25 (6896)	1.01 (5180)		GAC	0.83 (7557)	1.01 (3304)
	AAC	0.75 (4162)	0.99 (5066)	**Glu**	GAA	1.06 (13031)	1.22 (4382)
Lys	AAA	0.91 (8734)	1.02 (8793)		GAG*	0.94 (11441)	0.78 (2820)
	AAG*	1.09 (10508)	0.98 (8383)	**Gly**	GGU*	1.70 (6254)	0.69 (2104)
Cys	UGU*	1.22 (1402)	0.80 (7190)		GGC	1.03 (3784)	1.05 (3166)
	UGC	0.78 (897)	1.20 (10803)		GGA	0.97 (3577)	1.40 (4236)
Arg	CGU*	2.38 (5744)	0.63 (4026)		GGG	0.30 (1105)	0.86 (2612)
	CGC*	1.43 (3464)	0.78 (4958)	**Met**	AUG	1.00 (6682)	1.00 (7836)
	CGA*	1.24 (2988)	0.97 (6194)	**Trp**	UGG	1.00 (2063)	1.00 (10516)
	CGG	0.40 (965)	0.71 (4497)	**TER**	UAA	1.41 (356)	0.77 (5007)
	AGA	0.38 (919)	1.64 (10397)		UAG	0.91 (229)	0.64 (4149)
	AGG	0.18 (429)	1.27 (8063)		UGA	0.68 (171)	1.59 (10367)

Table 34 | Correspondence Analysis Results for the Stramenopile *P. halstedii*

Amino Acids	Codons	Highly Biased Genes	Weakly Biased Genes	Amino Acids	Codons	Highly Biased Genes	Weakly Biased Genes
Phe	UUU	0.78 (4965)	1.00 (3298)	**Ser**	UCU	0.35 (985)	0.73 (2697)
	UUC*	1.22 (7726)	1.00 (3274)		UCC*	1.42 (4046)	1.31 (4870)
Leu	UUA	0.05 (200)	0.35 (1330)		UCA	0.16 (464)	0.75 (2779)
	UUG*	1.87 (7811)	1.40 (5360)		UCG*	2.40 (6817)	1.56 (5775)
	CUU	0.32 (1340)	0.85 (3236)		AGU	0.33 (930)	0.69 (2563)
	CUC*	1.96 (8193)	1.46 (5562)		AGC*	1.34 (3819)	0.96 (3539)
	CUA	0.11 (443)	0.53 (2010)	**Thr**	ACU	0.24 (950)	0.70 (3233)
	CUG*	1.70 (7093)	1.42 (5421)		ACC*	1.45 (5721)	1.24 (5701)
Val	GUU	0.17 (900)	0.68 (3131)		ACA	0.24 (942)	0.85 (3921)
	GUC*	1.42 (7696)	1.25 (5721)		ACG*	2.08 (8218)	1.21 (5546)
	GUA	0.12 (657)	0.46 (2104)	**Tyr**	UAU	0.19 (868)	0.67 (1571)
	GUG*	2.29 (12386)	1.60 (7328)		UAC*	1.81 (8151)	1.33 (3146)
Pro	CCU	0.47 (1295)	0.83 (3590)	**Ala**	GCU	0.27 (1634)	0.74 (4501)
	CCC*	1.74 (4804)	1.09 (4730)		GCC*	1.81 (10936)	1.33 (8026)
	CCA	0.29 (789)	1.14 (4913)		GCA	0.24 (1428)	0.86 (5185)
	CCG*	1.51 (4179)	0.94 (4058)		GCG*	1.68 (10108)	1.08 (6514)
His	CAU	0.44 (1294)	0.84 (3743)	**Ile**	AUU	0.69 (3511)	1.03 (2935)
	CAC*	1.56 (4630)	1.16 (5183)		AUC*	2.28 (11624)	1.55 (4409)
Gln	CAA	0.97 (4401)	1.17 (6668)		AUA	0.03 (128)	0.41 (1169)
	CAG*	1.03 (4694)	0.83 (4690)	**Asp**	GAU	0.26 (1991)	0.64 (5290)
Asn	AAU	0.15 (837)	0.63 (2319)		GAC*	1.74 (13548)	1.36 (11249)
	AAC*	1.85 (10377)	1.37 (4986)	**Glu**	GAA	1.09 (8402)	1.09 (7009)
Lys	AAA	0.25 (2192)	0.77 (3763)		GAG	0.91 (6956)	0.91 (5854)
	AAG*	1.75 (15400)	1.23 (6011)	**Gly**	GGU	0.49 (2359)	0.85 (2651)
Cys	UGU	0.30 (592)	0.84 (1608)		GGC*	2.37 (11364)	1.46 (4552)
	UGC*	1.70 (3328)	1.16 (2216)		GGA	0.27 (1291)	0.87 (2705)
Arg	CGU	0.67 (1403)	1.03 (2706)		GGG*	0.87 (4185)	0.82 (2558)
	CGC*	3.86 (8022)	1.63 (4278)	**Met**	AUG	1.00 (6898)	1.00 (6062)
	CGA	0.52 (1082)	1.48 (3898)	**Trp**	UGG	1.00 (3461)	1.00 (2987)
	CGG	0.84 (1739)	1.05 (2757)	**TER**	UAA	1.91 (430)	1.11 (247)
	AGA	0.05 (110)	0.41 (1070)		UAG	0.86 (194)	0.92 (205)
	AGG	0.06 (117)	0.41 (1068)		UGA	0.23 (51)	0.97 (216)

Table 35 | Correspondence Analysis Results for the Stramenopile *A. astaci*

Amino Acids	Codons	Highly Biased Genes	Weakly Biased Genes	Amino Acids	Codons	Highly Biased Genes	Weakly Biased Genes
Phe	UUU	0.98 (6395)	0.95 (1653)	**Ser**	UCU	0.32 (894)	0.68 (1297)
	UUC	1.02 (6708)	1.05 (1834)		UCC	1.10 (3117)	1.13 (2147)
Leu	UUA	0.06 (226)	0.28 (696)		UCA	0.22 (623)	0.68 (1284)
	UUG	1.34 (5253)	1.64 (4103)		UCG*	2.70 (7634)	1.64 (3118)
	CUU	0.50 (1948)	0.77 (1935)		AGU	0.37 (1057)	0.67 (1273)
	CUC*	2.37 (9246)	1.54 (3859)		AGC*	1.29 (3661)	1.20 (2284)
	CUA	0.14 (540)	0.37 (916)	**Thr**	ACU	0.29 (1104)	0.66 (1453)
	CUG*	1.60 (6240)	1.41 (3521)		ACC	1.08 (4148)	1.07 (2357)
Val	GUU	0.36 (1966)	0.72 (1723)		ACA	0.40 (1543)	0.85 (1877)
	GUC*	1.81 (9762)	1.35 (3229)		ACG*	2.23 (8528)	1.41 (3097)
	GUA	0.14 (735)	0.40 (963)	**Tyr**	UAU	0.28 (1183)	0.62 (774)
	GUG*	1.69 (9107)	1.52 (3626)		UAC*	1.72 (7387)	1.38 (1703)
Pro	CCU	0.42 (1122)	0.92 (1241)	**Ala**	GCU	0.36 (1971)	0.74 (3348)
	CCC*	1.15 (3059)	0.92 (1244)		GCC*	1.45 (7976)	1.17 (5268)
	CCA	0.57 (1502)	1.11 (1507)		GCA	0.43 (2363)	0.92 (4122)
	CCG*	1.86 (4925)	1.05 (1424)		GCG*	1.77 (9722)	1.17 (5241)
His	CAU	0.49 (1290)	0.80 (1638)	**Ile**	AUU	0.85 (3421)	1.14 (2111)
	CAC*	1.51 (3978)	1.20 (2463)		AUC*	2.08 (8379)	1.54 (2833)
Gln	CAA	0.93 (3142)	1.11 (5673)		AUA	0.07 (290)	0.32 (590)
	CAG*	1.07 (3636)	0.89 (4503)	**Asp**	GAU	0.33 (1585)	0.62 (3806)
Asn	AAU	0.30 (1202)	0.64 (1680)		GAC*	1.67 (7886)	1.38 (8407)
	AAC*	1.70 (6760)	1.36 (3601)	**Glu**	GAA	0.94 (3167)	1.07 (8886)
Lys	AAA	0.45 (1561)	0.62 (4549)		GAG*	1.06 (3574)	0.93 (7702)
	AAG*	1.55 (5362)	1.38 (10047)	**Gly**	GGU	0.49 (1965)	0.89 (1368)
Cys	UGU	0.38 (915)	0.71 (521)		GGC*	2.01 (8037)	1.49 (2291)
	UGC*	1.62 (3964)	1.29 (938)		GGA	0.39 (1569)	0.85 (1306)
Arg	CGU	0.62 (930)	1.00 (1993)		GGG*	1.10 (4388)	0.76 (1166)
	CGC*	2.97 (4472)	1.82 (3614)	**Met**	AUG	1.00 (6686)	1.00 (4393)
	CGA	0.85 (1275)	1.38 (2746)	**Trp**	UGG	1.00 (4385)	1.00 (1273)
	CGG*	1.28 (1928)	1.03 (2040)	**TER**	UAA	1.20 (192)	1.23 (214)
	AGA	0.09 (139)	0.37 (730)		UAG	1.19 (191)	0.88 (152)
	AGG	0.19 (281)	0.41 (818)		UGA	0.61 (97)	0.89 (155)

Table 36 | Correspondence Analysis Results for the Stramenopile *A. invadans*

Amino Acids	Codons	Highly Biased Genes	Weakly Biased Genes	Amino Acids	Codons	Highly Biased Genes	Weakly Biased Genes
Phe	UUU	0.74 (4874)	1.15 (4130)	**Ser**	UCU	0.07 (197)	0.63 (2587)
	UUC*	1.26 (8367)	0.85 (3039)		UCC	0.76 (2292)	0.83 (3389)
Leu	UUA	0.00 (17)	0.36 (1754)		UCA	0.03 (91)	0.63 (2572)
	UUG	0.53 (2337)	1.03 (4967)		UCG*	3.68 (11143)	1.95 (7975)
	CUU	0.41 (1775)	0.98 (4693)		AGU	0.09 (279)	0.63 (2558)
	CUC*	4.41 (19284)	1.98 (9507)		AGC@	1.38 (4165)	1.32 (5405)
	CUA	0.01 (56)	0.37 (1763)	**Thr**	ACU	0.05 (229)	0.48 (2420)
	CUG	0.63 (2762)	1.28 (6172)		ACC@	1.00 (4311)	0.96 (4886)
Val	GUU	0.14 (776)	0.63 (3194)		ACA	0.07 (296)	0.81 (4103)
	GUC*	2.73 (15005)	1.65 (8404)		ACG*	2.88 (12396)	1.76 (8922)
	GUA	0.01 (79)	0.38 (1922)	**Tyr**	UAU	0.11 (532)	0.62 (1635)
	GUG	1.11 (6091)	1.34 (6797)		UAC*	1.89 (9278)	1.38 (3650)
Pro	CCU	0.13 (417)	0.58 (2622)	**Ala**	GCU	0.21 (1560)	0.65 (5343)
	CCC*	1.34 (4178)	0.81 (3674)		GCC*	1.98 (14524)	1.15 (9513)
	CCA	0.07 (222)	1.12 (5071)		GCA	0.08 (556)	0.81 (6662)
	CCG*	2.46 (7695)	1.50 (6789)		GCG*	1.73 (12703)	1.39 (11482)
His	CAU	0.19 (649)	0.73 (3051)	**Ile**	AUU	0.66 (3623)	1.06 (3365)
	CAC*	1.81 (6155)	1.27 (5275)		AUC*	2.34 (12867)	1.55 (4914)
Gln	CAA	0.27 (1405)	1.01 (5033)		AUA	0.00 (9)	0.38 (1216)
	CAG*	1.73 (8944)	0.99 (4885)	**Asp**	GAU	0.21 (1916)	0.58 (4181)
Asn	AAU	0.04 (270)	0.69 (2415)		GAC*	1.79 (16047)	1.42 (10172)
	AAC*	1.96 (12227)	1.31 (4567)	**Glu**	GAA	0.50 (4826)	0.71 (3998)
Lys	AAA	0.04 (426)	0.72 (2858)		GAG*	1.50 (14322)	1.29 (7331)
	AAG*	1.96 (22412)	1.28 (5055)	**Gly**	GGU	0.60 (3191)	1.01 (3988)
Cys	UGU	0.07 (156)	0.58 (1705)		GGC*	3.21 (17077)	1.82 (7161)
	UGC*	1.93 (4363)	1.42 (4135)		GGA	0.02 (85)	0.57 (2263)
Arg	CGU	0.55 (1294)	1.00 (3527)		GGG	0.17 (910)	0.60 (2368)
	CGC*	4.93 (11689)	2.11 (7418)	**Met**	AUG	1.00 (8788)	1.00 (5512)
	CGA	0.07 (158)	1.17 (4131)	**Trp**	UGG	1.00 (3629)	1.00 (3631)
	CGG	0.44 (1035)	1.05 (3700)	**TER**	UAA	2.26 (719)	0.95 (313)
	AGA	0.01 (29)	0.38 (1329)		UAG	0.65 (207)	0.86 (284)
	AGG	0.01 (13)	0.29 (1015)		UGA	0.09 (29)	1.19 (395)

Table 37 | Correspondence Analysis Results for the Stramenopile *S. parasitica*

Amino Acids	Codons	Highly Biased Genes	Weakly Biased Genes	Amino Acids	Codons	Highly Biased Genes	Weakly Biased Genes
Phe	UUU	1.00 (3947)	1.20 (8516)	**Ser**	UCU	0.18 (630)	1.32 (9368)
	UUC*	1.00 (3947)	0.80 (5637)		UCC*	2.16 (7725)	0.68 (4800)
Leu	UUA	0.01 (21)	0.85 (4956)		UCA	0.04 (127)	1.35 (9550)
	UUG	0.20 (819)	1.64 (9537)		UCG*	1.74 (6208)	0.96 (6814)
	CUU	0.24 (992)	1.31 (7647)		AGU	0.06 (221)	0.98 (6935)
	CUC*	2.70 (11247)	0.76 (4457)		AGC*	1.83 (6554)	0.70 (4983)
	CUA	0.06 (262)	0.73 (4273)	**Thr**	ACU	0.06 (201)	1.17 (6841)
	CUG*	2.79 (11614)	0.71 (4111)		ACC*	2.34 (8530)	0.69 (4073)
Val	GUU	0.27 (1170)	1.42 (8161)		ACA	0.20 (725)	1.37 (8025)
	GUC*	2.39 (10165)	0.85 (4888)		ACG*	1.41 (5141)	0.77 (4509)
	GUA	0.07 (302)	0.89 (5120)	**Tyr**	UAU	0.12 (355)	1.09 (5031)
	GUG*	1.27 (5393)	0.84 (4844)		UAC*	1.88 (5424)	0.91 (4183)
Pro	CCU	0.16 (683)	1.20 (5293)	**Ala**	GCU	0.16 (1232)	1.26 (8304)
	CCC*	1.89 (8163)	0.59 (2622)		GCC*	2.36 (17611)	0.68 (4498)
	CCA	0.24 (1030)	1.52 (6732)		GCA	0.28 (2082)	1.45 (9557)
	CCG*	1.72 (7437)	0.69 (3046)		GCG*	1.20 (8964)	0.61 (4022)
His	CAU	0.11 (336)	1.20 (5126)	**Ile**	AUU	0.34 (870)	1.36 (10122)
	CAC*	1.89 (5967)	0.80 (3408)		AUC*	2.58 (6699)	0.89 (6605)
Gln	CAA	0.37 (1433)	1.29 (10142)		AUA	0.08 (215)	0.75 (5614)
	CAG*	1.63 (6338)	0.71 (5527)	**Asp**	GAU	0.22 (1795)	1.27 (16831)
Asn	AAU	0.13 (498)	1.21 (12467)		GAC*	1.78 (14188)	0.73 (9661)
	AAC*	1.87 (7424)	0.79 (8118)	**Glu**	GAA	0.53 (4789)	1.35 (21103)
Lys	AAA	0.33 (1642)	1.19 (15812)		GAG*	1.47 (13178)	0.65 (10101)
	AAG*	1.67 (8218)	0.81 (10664)	**Gly**	GGU	0.19 (1291)	1.18 (6707)
Cys	UGU	0.14 (264)	1.08 (3279)		GGC*	1.65 (10952)	0.68 (3884)
	UGC*	1.86 (3581)	0.92 (2819)		GGA	0.60 (3991)	1.50 (8521)
Arg	CGU	0.14 (500)	0.92 (3611)		GGG*	1.55 (10279)	0.64 (3624)
	CGC*	1.47 (5066)	0.62 (2442)	**Met**	AUG	1.00 (4786)	1.00 (9252)
	CGA	0.33 (1133)	1.42 (5595)	**Trp**	UGG	1.00 (3349)	1.00 (3964)
	CGG*	2.76 (9527)	0.54 (2112)	**TER**	UAA	0.39 (77)	1.15 (228)
	AGA	0.17 (572)	1.70 (6680)		UAG	1.48 (294)	0.72 (144)
	AGG*	1.14 (3937)	0.80 (3145)		UGA	1.13 (225)	1.13 (225)

Table 38 | Correspondence Analysis Results for the Stramenopile *P. multistriata*

Amino Acids	Codons	Highly Biased Genes	Weakly Biased Genes	Amino Acids	Codons	Highly Biased Genes	Weakly Biased Genes
Phe	UUU	0.72 (2522)	1.12 (4000)	**Ser**	UCU*	1.37 (2940)	1.12 (4510)
	UUC*	1.28 (4520)	0.88 (3152)		UCC*	1.86 (3977)	0.65 (2606)
Leu	UUA	0.08 (205)	0.95 (3444)		UCA	0.43 (914)	1.61 (6484)
	UUG*	1.96 (4926)	1.61 (5792)		UCG	0.70 (1499)	0.97 (3926)
	CUU*	1.53 (3863)	0.93 (3373)		AGU*	1.07 (2282)	0.95 (3816)
	CUC*	2.03 (5109)	0.80 (2879)		AGC	0.57 (1223)	0.71 (2863)
	CUA	0.19 (487)	0.68 (2456)	**Thr**	ACU*	1.20 (3336)	0.99 (3367)
	CUG	0.20 (515)	1.03 (3702)		ACC*	1.81 (5022)	0.51 (1744)
Val	GUU*	1.23 (4036)	1.15 (3780)		ACA	0.34 (955)	1.55 (5281)
	GUC*	1.33 (4376)	0.71 (2332)		ACG	0.65 (1801)	0.94 (3199)
	GUA	0.29 (956)	1.00 (3280)	**Tyr**	UAU	0.52 (1449)	1.08 (2942)
	GUG	1.15 (3774)	1.14 (3729)		UAC*	1.48 (4129)	0.92 (2515)
Pro	CCU*	1.39 (2826)	0.98 (2932)	**Ala**	GCU*	1.35 (6593)	1.12 (3389)
	CCC*	1.68 (3407)	0.44 (1324)		GCC*	1.71 (8356)	0.52 (1584)
	CCA	0.63 (1282)	1.70 (5097)		GCA	0.54 (2632)	1.79 (5414)
	CCG	0.30 (618)	0.88 (2653)		GCG	0.39 (1925)	0.57 (1741)
His	CAU	0.76 (1191)	1.17 (3669)	**Ile**	AUU*	1.45 (4723)	1.16 (4092)
	CAC*	1.24 (1941)	0.83 (2611)		AUC*	1.49 (4842)	0.93 (3296)
Gln	CAA	0.84 (2585)	1.40 (6383)		AUA	0.05 (178)	0.91 (3198)
	CAG*	1.16 (3566)	0.60 (2720)	**Asp**	GAU	1.19 (6939)	1.24 (6316)
Asn	AAU	0.69 (2659)	1.08 (4533)		GAC*	0.81 (4721)	0.76 (3896)
	AAC*	1.31 (4993)	0.92 (3893)	**Glu**	GAA	0.64 (4408)	1.17 (6339)
Lys	AAA	0.41 (2770)	1.10 (4539)		GAG*	1.36 (9313)	0.83 (4512)
	AAG*	1.59 (10775)	0.90 (3717)	**Gly**	GGU	0.89 (3364)	1.04 (3406)
Cys	UGU	0.92 (1434)	1.02 (2942)		GGC	0.55 (2084)	0.71 (2310)
	UGC*	1.08 (1676)	0.98 (2828)		GGA*	2.16 (8127)	1.50 (4905)
Arg	CGU*	2.89 (4075)	0.92 (2515)		GGG	0.39 (1470)	0.74 (2425)
	CGC*	1.03 (1448)	0.42 (1163)	**Met**	AUG	1.00 (5493)	1.00 (5205)
	CGA	0.38 (533)	1.24 (3387)	**Trp**	UGG	1.00 (1956)	1.00 (4323)
	CGG	0.11 (152)	0.70 (1918)	**TER**	UAA	2.03 (345)	0.51 (687)
	AGA	0.35 (497)	1.69 (4632)		UAG	0.62 (105)	0.53 (710)
	AGG*	1.24 (1748)	1.03 (2829)		UGA	0.35 (59)	1.96 (2645)

Table 39 | Correspondence Analysis Results for the Stramenopile *T. pseudonana*

Amino Acids	Codons	Highly Biased Genes	Weakly Biased Genes	Amino Acids	Codons	Highly Biased Genes	Weakly Biased Genes
Phe	UUU	0.47 (1784)	1.38 (3744)	**Ser**	UCU	0.55 (1435)	1.39 (2074)
	UUC*	1.53 (5832)	0.62 (1691)		UCC*	2.44 (6363)	0.64 (951)
Leu	UUA	0.09 (335)	1.70 (3277)		UCA	0.26 (666)	1.34 (2000)
	UUG	1.09 (4176)	1.19 (2286)		UCG*	1.12 (2928)	0.75 (1114)
	CUU	0.49 (1866)	1.27 (2437)		AGU	0.32 (828)	1.12 (1673)
	CUC*	2.15 (8282)	0.49 (939)		AGC*	1.31 (3407)	0.75 (1121)
	CUA	0.16 (600)	0.71 (1361)	**Thr**	ACU	0.34 (889)	1.15 (1799)
	CUG*	2.03 (7823)	0.65 (1248)		ACC*	1.74 (4576)	0.60 (934)
Val	GUU	0.25 (970)	1.33 (2251)		ACA	0.35 (926)	1.55 (2418)
	GUC*	1.41 (5469)	0.72 (1229)		ACG*	1.56 (4105)	0.70 (1091)
	GUA	0.19 (740)	1.14 (1930)	**Tyr**	UAU	0.30 (732)	1.27 (2590)
	GUG*	2.15 (8382)	0.82 (1385)		UAC*	1.70 (4215)	0.73 (1486)
Pro	CCU	0.77 (2749)	1.47 (1665)	**Ala**	GCU	0.37 (2065)	1.33 (2173)
	CCC*	1.96 (7015)	0.62 (708)		GCC*	1.92 (10657)	0.63 (1020)
	CCA	0.29 (1035)	1.31 (1487)		GCA	0.36 (1990)	1.40 (2277)
	CCG*	0.98 (3524)	0.60 (682)		GCG*	1.35 (7489)	0.65 (1058)
His	CAU	0.44 (1062)	1.32 (2197)	**Ile**	AUU	0.51 (1199)	1.46 (3536)
	CAC*	1.56 (3715)	0.68 (1126)		AUC*	2.30 (5366)	0.66 (1607)
Gln	CAA	0.50 (1804)	1.40 (3040)		AUA	0.19 (449)	0.88 (2141)
	CAG*	1.50 (5363)	0.60 (1314)	**Asp**	GAU	0.39 (1977)	1.33 (3322)
Asn	AAU	0.44 (1092)	1.35 (3234)		GAC*	1.61 (8283)	0.67 (1659)
	AAC*	1.56 (3835)	0.65 (1545)	**Glu**	GAA	0.40 (3460)	1.29 (3531)
Lys	AAA	0.44 (2123)	1.29 (4434)		GAG*	1.60 (13722)	0.71 (1935)
	AAG*	1.56 (7631)	0.71 (2426)	**Gly**	GGU	0.34 (1815)	1.21 (2127)
Cys	UGU	0.43 (574)	1.12 (1515)		GGC*	1.40 (7364)	0.69 (1216)
	UGC*	1.57 (2109)	0.88 (1188)		GGA	0.58 (3080)	1.45 (2537)
Arg	CGU	0.52 (1348)	1.16 (1501)		GGG*	1.67 (8831)	0.65 (1140)
	CGC*	1.68 (4322)	0.62 (804)	**Met**	AUG	1.00 (5184)	1.00 (2859)
	CGA	0.59 (1534)	1.15 (1486)	**Trp**	UGG	1.00 (2761)	1.00 (1787)
	CGG*	2.04 (5253)	0.54 (698)	**TER**	UAA	0.65 (110)	1.23 (821)
	AGA	0.26 (659)	1.59 (2059)		UAG	0.90 (152)	0.82 (549)
	AGG	0.91 (2354)	0.93 (1204)		UGA	1.45 (246)	0.94 (628)

Table 40 | Correspondence Analysis Results for the Stramenopile *N. gaditana*

Amino Acids	Codons	Highly Biased Genes	Weakly Biased Genes	Amino Acids	Codons	Highly Biased Genes	Weakly Biased Genes
Phe	UUU	0.02 (42)	1.14 (5826)	**Ser**	UCU	0.00 (6)	1.26 (4725)
	UUC*	1.98 (5474)	0.86 (4438)		UCC*	2.21 (3019)	0.85 (3182)
Leu	UUA	0.00 (11)	0.53 (2652)		UCA	0.00 (6)	1.28 (4806)
	UUG	0.05 (124)	1.05 (5282)		UCG*	2.33 (3189)	0.92 (3469)
	CUU	0.01 (13)	1.46 (7368)		AGU	0.01 (8)	0.83 (3133)
	CUC*	4.78 (12062)	1.21 (6081)		AGC*	1.44 (1975)	0.87 (3269)
	CUA	0.00 (7)	0.76 (3832)	**Thr**	ACU	0.00 (2)	1.11 (4271)
	CUG*	1.15 (2911)	1.00 (5023)		ACC	0.78 (1402)	0.83 (3196)
Val	GUU	0.00 (10)	1.15 (5467)		ACA	0.00 (4)	1.24 (4766)
	GUC*	3.01 (9358)	1.10 (5207)		ACG*	3.22 (5818)	0.81 (3094)
	GUA	0.00 (5)	0.67 (3195)	**Tyr**	UAU	0.00 (7)	1.06 (3682)
	GUG	0.99 (3065)	1.07 (5071)		UAC*	2.00 (3827)	0.94 (3237)
Pro	CCU	0.01 (12)	1.11 (3963)	**Ala**	GCU	0.00 (26)	1.11 (7817)
	CCC*	1.80 (4273)	0.78 (2796)		GCC*	1.51 (10770)	0.84 (5929)
	CCA	0.01 (16)	1.31 (4700)		GCA	0.00 (26)	1.25 (8821)
	CCG*	2.18 (5172)	0.79 (2841)		GCG*	2.48 (17699)	0.79 (5570)
His	CAU	0.00 (7)	1.09 (4993)	**Ile**	AUU	0.01 (7)	1.24 (5801)
	CAC*	2.00 (3106)	0.91 (4206)		AUC*	2.99 (4048)	1.03 (4789)
Gln	CAA	0.01 (25)	1.09 (6255)		AUA	0.00 (2)	0.73 (3418)
	CAG*	1.99 (3318)	0.91 (5255)	**Asp**	GAU	0.01 (48)	1.17 (9480)
Asn	AAU	0.01 (14)	1.18 (5934)		GAC*	1.99 (11651)	0.83 (6790)
	AAC*	1.99 (3304)	0.82 (4118)	**Glu**	GAA	0.02 (81)	1.07 (8998)
Lys	AAA	0.02 (73)	1.02 (6221)		GAG*	1.98 (8942)	0.93 (7796)
	AAG*	1.98 (6200)	0.98 (5987)	**Gly**	GGU	0.01 (44)	1.10 (5272)
Cys	UGU	0.00 (7)	1.00 (3198)		GGC*	3.64 (13764)	1.16 (5570)
	UGC*	2.00 (2917)	1.00 (3220)		GGA	0.01 (36)	1.00 (4792)
Arg	CGU	0.01 (16)	1.24 (4811)		GGG	0.34 (1282)	0.75 (3622)
	CGC*	4.03 (7791)	1.46 (5685)	**Met**	AUG	1.00 (2759)	1.00 (6155)
	CGA	0.02 (38)	1.27 (4936)	**Trp**	UGG	1.00 (2098)	1.00 (3443)
	CGG*	1.88 (3625)	0.81 (3150)	**TER**	UAA	0.64 (57)	0.81 (132)
	AGA	0.00 (4)	0.69 (2702)		UAG	1.53 (137)	0.92 (149)
	AGG	0.06 (119)	0.53 (2067)		UGA	0.84 (75)	1.27 (207)

Table 41 | Correspondence Analysis Results for the Stramenopile *A. anophagefferens*

Amino Acids	Codons	Highly Biased Genes	Weakly Biased Genes	Amino Acids	Codons	Highly Biased Genes	Weakly Biased Genes
Phe	UUU	0.51 (5641)	1.19 (9240)	**Ser**	UCU	0.69 (5370)	1.31 (8564)
	UUC*	1.49 (16463)	0.81 (6247)		UCC*	1.41 (10984)	0.64 (4174)
Leu	UUA	0.18 (1608)	1.22 (7148)		UCA	0.59 (4549)	1.45 (9461)
	UUG	0.68 (5947)	1.28 (7491)		UCG*	1.31 (10175)	0.57 (3727)
	CUU	0.80 (7074)	1.34 (7861)		AGU	0.43 (3305)	1.21 (7914)
	CUC*	2.12 (18645)	0.70 (4089)		AGC*	1.58 (12270)	0.82 (5392)
	CUA	0.42 (3719)	0.82 (4811)	**Thr**	ACU	0.54 (4013)	1.26 (7612)
	CUG*	1.80 (15828)	0.65 (3816)		ACC*	1.28 (9511)	0.67 (4041)
Val	GUU	0.64 (5933)	1.39 (8087)		ACA	0.67 (4975)	1.52 (9166)
	GUC*	1.30 (12053)	0.66 (3809)		ACG*	1.51 (11207)	0.55 (3348)
	GUA	0.39 (3619)	1.03 (5971)	**Tyr**	UAU	0.49 (3845)	1.17 (6782)
	GUG*	1.67 (15530)	0.93 (5388)		UAC*	1.51 (11916)	0.83 (4768)
Pro	CCU	0.83 (5527)	1.42 (6010)	**Ala**	GCU	0.77 (9589)	1.42 (9267)
	CCC*	1.25 (8342)	0.51 (2175)		GCC*	1.18 (14726)	0.49 (3235)
	CCA	0.74 (4926)	1.58 (6698)		GCA	0.74 (9217)	1.64 (10710)
	CCG*	1.18 (7874)	0.49 (2054)		GCG*	1.30 (16195)	0.45 (2934)
His	CAU	0.59 (4063)	1.29 (6750)	**Ile**	AUU	0.70 (6164)	1.29 (10833)
	CAC*	1.41 (9651)	0.71 (3693)		AUC*	1.87 (16567)	0.74 (6181)
Gln	CAA	0.58 (6157)	1.29 (11186)		AUA	0.44 (3876)	0.97 (8178)
	CAG*	1.42 (15241)	0.71 (6183)	**Asp**	GAU	0.73 (11215)	1.41 (17968)
Asn	AAU	0.55 (5789)	1.30 (15619)		GAC*	1.27 (19426)	0.59 (7519)
	AAC*	1.45 (15109)	0.70 (8438)	**Glu**	GAA	0.53 (9709)	1.28 (19818)
Lys	AAA	0.52 (8941)	1.20 (22386)		GAG*	1.47 (27054)	0.72 (11236)
	AAG*	1.48 (25400)	0.80 (14881)	**Gly**	GGU	0.48 (5165)	1.01 (6179)
Cys	UGU	0.42 (1911)	1.13 (3725)		GGC*	1.46 (15620)	0.57 (3489)
	UGC*	1.58 (7185)	0.87 (2877)		GGA	0.82 (8746)	1.76 (10758)
Arg	CGU	0.59 (3551)	0.81 (3540)		GGG*	1.24 (13227)	0.67 (4091)
	CGC*	1.74 (10568)	0.41 (1808)	**Met**	AUG	1.00 (14824)	1.00 (10522)
	CGA	0.79 (4804)	1.04 (4557)	**Trp**	UGG	1.00 (7378)	1.00 (4226)
	CGG*	1.09 (6637)	0.43 (1895)	**TER**	UAA	0.81 (235)	1.18 (407)
	AGA	0.64 (3914)	2.22 (9686)		UAG	0.99 (289)	0.70 (241)
	AGG*	1.14 (6945)	1.08 (4739)		UGA	1.20 (350)	1.13 (391)

Table 42 | Correspondence Analysis Results for the Rhizarian *B. natans*

Amino Acids	Codons	Highly Biased Genes	Weakly Biased Genes	Amino Acids	Codons	Highly Biased Genes	Weakly Biased Genes
Phe	UUU	0.08 (374)	0.75 (4669)	**Ser**	UCU	0.06 (140)	1.15 (6223)
	UUC*	1.92 (8610)	1.25 (7773)		UCC*	1.21 (3046)	1.01 (5488)
Leu	UUA	0.01 (55)	0.38 (2152)		UCA	0.08 (203)	0.94 (5075)
	UUG	0.23 (848)	1.08 (6134)		UCG*	3.13 (7861)	1.53 (8287)
	CUU	0.08 (288)	1.34 (7624)		AGU	0.06 (161)	0.58 (3124)
	CUC*	1.19 (4410)	1.00 (5718)		AGC*	1.46 (3665)	0.79 (4270)
	CUA	0.06 (227)	0.51 (2881)	**Thr**	ACU	0.04 (121)	0.92 (4169)
	CUG*	4.43 (16488)	1.70 (9703)		ACC*	1.24 (4034)	0.91 (4106)
Val	GUU	0.14 (681)	1.32 (8454)		ACA	0.10 (320)	0.87 (3915)
	GUC*	2.65 (12767)	1.36 (8703)		ACG*	2.62 (8517)	1.30 (5843)
	GUA	0.04 (189)	0.45 (2876)	**Tyr**	UAU	0.07 (221)	0.87 (2766)
	GUG*	1.17 (5658)	0.88 (5646)		UAC*	1.93 (5686)	1.13 (3598)
Pro	CCU	0.12 (350)	1.21 (5588)	**Ala**	GCU	0.08 (557)	1.10 (8459)
	CCC*	1.18 (3357)	0.69 (3213)		GCC*	2.09 (13827)	0.98 (7525)
	CCA	0.16 (441)	0.99 (4574)		GCA	0.10 (630)	0.90 (6867)
	CCG*	2.54 (7196)	1.11 (5122)		GCG*	1.73 (11498)	1.02 (7836)
His	CAU	0.26 (698)	1.10 (4791)	**Ile**	AUU	0.18 (599)	1.22 (6647)
	CAC*	1.74 (4696)	0.90 (3933)		AUC*	2.74 (8960)	1.33 (7255)
Gln	CAA	0.13 (561)	0.87 (5879)		AUA	0.08 (247)	0.45 (2461)
	CAG*	1.87 (8332)	1.13 (7594)	**Asp**	GAU	0.17 (1315)	0.98 (9930)
Asn	AAU	0.08 (277)	0.86 (4529)		GAC*	1.83 (14078)	1.02 (10330)
	AAC*	1.92 (6661)	1.14 (6002)	**Glu**	GAA	0.27 (1310)	1.04 (9216)
Lys	AAA	0.17 (666)	0.79 (4617)		GAG*	1.73 (8550)	0.96 (8572)
	AAG*	1.83 (7154)	1.21 (7017)	**Gly**	GGU	0.17 (633)	0.97 (4444)
Cys	UGU	0.14 (258)	0.81 (2812)		GGC*	2.71 (10185)	1.34 (6138)
	UGC*	1.86 (3478)	1.19 (4133)		GGA	0.08 (298)	0.87 (3973)
Arg	CGU	0.32 (859)	1.32 (5495)		GGG*	1.04 (3895)	0.82 (3754)
	CGC*	3.04 (8177)	1.32 (5475)	**Met**	AUG	1.00 (5639)	1.00 (6827)
	CGA	0.33 (894)	1.21 (5005)	**Trp**	UGG	1.00 (2949)	1.00 (3885)
	CGG*	1.79 (4809)	1.05 (4362)	**TER**	UAA	0.71 (105)	0.96 (153)
	AGA	0.03 (84)	0.53 (2183)		UAG	0.81 (120)	0.63 (100)
	AGG	0.49 (1332)	0.58 (2400)		UGA	1.48 (220)	1.41 (224)

Table 42 | Correspondence Analysis Results for the Rhizarian *P. brassicae*

6.3 Codon Usage – Expression Data

General description of tables 43-75: The codon usage data is given in two forms, inside and outside of brackets. The numbers inside of the brackets represent the amount the codon was found within the genome files. The numbers outside the brackets represent the proportion of that the codon was found within the set of genes as opposed to synonymous codons. The p-values shown were calculated by a 2x2 contingency table, any p-value above 0.001 was deemed insignificant, these are shown using red text.

Amino Acids	Codons	Top 5% Expressed Genes	Bottom 5% Expressed Genes	Chi-Sqaured	Amino Acids	Codons	Top 5% Expressed Genes	Bottom 5% Expressed Genes	Chi-Sqaured
Phe	UUU	0.10 (1036)	0.16 (1578)	p<0.0001	Ser	UCU	0.08 (1194)	0.07 (1648)	p=0.0023
	UUC*	0.90 (9269)	0.84 (8149)	p<0.0001		UCC*	0.29 (4568)	0.21 (5039)	p<0.0001
Leu	UUA	0.00 (33)	0.01 (177)	p<0.0001		UCA	0.02 (363)	0.06 (1556)	p<0.0001
	UUG	0.06 (1309)	0.09 (2309)	p<0.0001		UCG	0.34 (5343)	0.33 (7939)	p=0.0089
	CUU*	0.08 (1768)	0.07 (1750)	p<0.0001		AGU	0.02 (296)	0.04 (1033)	p<0.0001
	CUC*	0.54 (12023)	0.41 (10710)	p<0.0001		AGC	0.25 (3865)	0.29 (6901)	p<0.0001
	CUA	0.01 (145)	0.03 (739)	p<0.0001	Thr	ACU*	0.10 (1490)	0.08 (1523)	p<0.0001
	CUG	0.32 (7046)	0.40 (10480)	p<0.0001		ACC*	0.63 (9036)	0.53 (9564)	p<0.0001
Val	GUU*	0.08 (1430)	0.07 (1386)	p<0.0001		ACA	0.03 (467)	0.08 (1531)	p<0.0001
	GUC*	0.46 (8272)	0.40 (8168)	p<0.0001		ACG	0.23 (3244)	0.30 (5488)	p<0.0001
	GUA	0.02 (359)	0.04 (780)	p<0.0001	Tyr	UAU	0.09 (708)	0.14 (953)	p<0.0001
	GUG	0.44 (8024)	0.50 (10186)	p<0.0001		UAC*	0.91 (6920)	0.86 (5960)	p<0.0001
Pro	CCU	0.12 (1443)	0.11 (1860)	p=0.0256	Ala	GCU*	0.15 (3370)	0.10 (2914)	p<0.0001
	CCC*	0.60 (7196)	0.40 (6646)	p<0.0001		GCC*	0.65 (14570)	0.54 (16217)	p<0.0001
	CCA	0.03 (372)	0.12 (1939)	p<0.0001		GCA	0.03 (651)	0.08 (2432)	p<0.0001
	CCG	0.24 (2917)	0.37 (6104)	p<0.0001		GCG	0.17 (3862)	0.28 (8322)	p<0.0001
His	CAU	0.10 (599)	0.15 (1203)	p<0.0001	Ile	AUU	0.14 (1759)	0.13 (1377)	p=0.1825
	CAC*	0.90 (5403)	0.85 (6596)	p<0.0001		AUC*	0.85 (10978)	0.81 (8543)	p<0.0001
Gln	CAA	0.10 (899)	0.21 (2308)	p<0.0001		AUA	0.01 (127)	0.06 (610)	p<0.0001
	CAG*	0.90 (8355)	0.79 (8752)	p<0.0001	Asp	GAU*	0.24 (3490)	0.20 (3103)	p<0.0001
Asn	AAU	0.09 (822)	0.15 (1455)	p<0.0001		GAC	0.76 (10905)	0.80 (12326)	p<0.0001
	AAC*	0.91 (8732)	0.85 (8220)	p<0.0001	Glu	GAA	0.12 (2129)	0.20 (3377)	p<0.0001
Lys	AAA	0.03 (608)	0.11 (1248)	p<0.0001		GAG*	0.88 (15130)	0.80 (13105)	p<0.0001
	AAG*	0.97 (16910)	0.89 (10115)	p<0.0001	Gly	GGU*	0.23 (4069)	0.11 (2373)	p<0.0001
Cys	UGU	0.10 (315)	0.19 (918)	p<0.0001		GGC	0.65 (11292)	0.63 (13212)	p=0.0029
	UGC*	0.90 (2843)	0.81 (4044)	p<0.0001		GGA	0.08 (1334)	0.10 (2187)	p<0.0001
Arg	CGU*	0.13 (1724)	0.07 (1230)	p<0.0001		GGG	0.04 (736)	0.15 (3095)	p<0.0001
	CGC*	0.45 (5797)	0.42 (7122)	p<0.0001					
	CGA	0.07 (913)	0.12 (1985)	p<0.0001					
	CGG	0.08 (1029)	0.22 (3777)	p<0.0001					
	AGA	0.04 (466)	0.05 (817)	p<0.0001					
	AGG*	0.22 (2868)	0.13 (2198)	p<0.0001					

Table 43 | Codon Usage from Expression Data for the Amoebozoan *A. castellani*

Amino Acids	Codons	Top 5% Expressed Genes	Bottom 5% Expressed Genes	Chi-Sqaured
Phe	UUU	0.52 (1768)	0.85 (6455)	p<0.0001
	UUC*	0.48 (1660)	0.15 (1113)	p<0.0001
Leu	UUA	0.33 (2238)	0.60 (7594)	p<0.0001
	UUG	0.07 (476)	0.08 (1056)	p=0.0004
	CUU*	0.53 (3647)	0.21 (2695)	p<0.0001
	CUC*	0.06 (420)	0.03 (417)	p<0.0001
	CUA	0.01 (62)	0.06 (743)	p<0.0001
	CUG	0.00 (22)	0.01 (117)	p<0.0001
Val	GUU*	0.70 (4933)	0.52 (3414)	p<0.0001
	GUC*	0.12 (874)	0.06 (415)	p<0.0001
	GUA	0.16 (1134)	0.35 (2321)	p<0.0001
	GUG	0.01 (91)	0.07 (438)	p<0.0001
Pro	CCU	0.08 (311)	0.32 (1492)	p<0.0001
	CCC	0.01 (28)	0.03 (163)	p<0.0001
	CCA*	0.91 (3447)	0.63 (2976)	p<0.0001
	CCG	0.00 (4)	0.02 (99)	p<0.0001
His	CAU	0.75 (1463)	0.91 (2289)	p<0.0001
	CAC*	0.25 (488)	0.09 (231)	p<0.0001
Gln	CAA*	0.98 (3273)	0.92 (4974)	p<0.0001
	CAG	0.02 (60)	0.08 (434)	p<0.0001
Asn	AAU	0.75 (3596)	0.89 (11076)	p<0.0001
	AAC*	0.25 (1176)	0.11 (1382)	p<0.0001
Lys	AAA	0.65 (7106)	0.85 (10702)	p<0.0001
	AAG*	0.35 (3854)	0.15 (1927)	p<0.0001
Cys	UGU	0.78 (1900)	0.94 (3488)	p<0.0001
	UGC*	0.22 (534)	0.06 (205)	p<0.0001
Arg	CGU	0.10 (448)	0.12 (581)	p<0.0001
	CGC	0.00 (13)	0.01 (59)	p<0.0001
	CGA	0.02 (78)	0.10 (489)	p<0.0001
	CGG	0.00 (1)	0.01 (50)	p<0.0001
	AGA*	0.87 (4023)	0.70 (3300)	p<0.0001
	AGG	0.01 (40)	0.05 (255)	p<0.0001
Ser	UCU	0.24 (1274)	0.26 (2974)	p=0.0002
	UCC*	0.04 (216)	0.02 (254)	p<0.0001
	UCA*	0.50 (2687)	0.42 (4782)	p<0.0001
	UCG	0.01 (47)	0.02 (206)	p<0.0001
	AGU	0.18 (1000)	0.26 (2955)	p<0.0001
	AGC*	0.04 (195)	0.02 (188)	p<0.0001
Thr	ACU*	0.55 (3144)	0.34 (2934)	p<0.0001
	ACC*	0.10 (572)	0.04 (321)	p<0.0001
	ACA	0.34 (1929)	0.59 (5023)	p<0.0001
	ACG	0.01 (71)	0.03 (251)	p<0.0001
Tyr	UAU	0.79 (2502)	0.91 (6691)	p<0.0001
	UAC*	0.21 (668)	0.09 (653)	p<0.0001
Ala	GCU*	0.61 (4885)	0.40 (1354)	p<0.0001
	GCC*	0.09 (682)	0.04 (134)	p<0.0001
	GCA	0.30 (2397)	0.53 (1809)	p<0.0001
	GCG	0.00 (18)	0.03 (103)	p<0.0001
Ile	AUU*	0.83 (4945)	0.58 (9605)	p<0.0001
	AUC*	0.12 (719)	0.06 (945)	p<0.0001
	AUA	0.05 (269)	0.36 (5948)	p<0.0001
Asp	GAU	0.83 (4565)	0.88 (6254)	p<0.0001
	GAC*	0.17 (936)	0.12 (891)	p<0.0001
Glu	GAA*	0.94 (6427)	0.90 (9725)	p<0.0001
	GAG	0.06 (443)	0.10 (1116)	p<0.0001
Gly	GGU	0.29 (1962)	0.27 (1608)	p=0.0021
	GGC	0.01 (82)	0.02 (129)	p<0.0001
	GGA*	0.69 (4603)	0.64 (3847)	p<0.0001
	GGG	0.01 (67)	0.07 (424)	p<0.0001

Table 44 | Codon Usage from Expression Data for the Amoebozoan *E. histolytica*

Amino Acids	Codons	Top 5% Expressed Genes	Bottom 5% Expressed Genes	Chi-Sqaured	Amino Acids	Codons	Top 5% Expressed Genes	Bottom 5% Expressed Genes	Chi-Sqaured
Phe	UUU	0.53 (3892)	0.78 (14984)	p<0.0001	**Ser**	UCU*	0.20 (3010)	0.15 (4693)	p<0.0001
	UUC*	0.47 (3484)	0.22 (4114)	p<0.0001		UCC*	0.08 (1173)	0.03 (974)	p<0.0001
Leu	UUA	0.58 (7367)	0.73 (21898)	p<0.0001		UCA	0.51 (7822)	0.56 (16873)	p<0.0001
	UUG	0.10 (1326)	0.10 (2950)	p=0.0266		UCG	0.01 (199)	0.02 (622)	p<0.0001
	CUU*	0.13 (1679)	0.10 (3030)	p<0.0001		AGU	0.15 (2325)	0.22 (6728)	p<0.0001
	CUC*	0.15 (1957)	0.02 (519)	p<0.0001		AGC*	0.05 (734)	0.02 (498)	p<0.0001
	CUA	0.02 (298)	0.06 (1673)	p<0.0001	**Thr**	ACU*	0.40 (5111)	0.35 (5939)	p<0.0001
	CUG	0.00 (24)	0.00 (102)	p=0.0097		ACC*	0.29 (3703)	0.09 (1619)	p<0.0001
Val	GUU	0.57 (5965)	0.57 (7411)	p=0.7463		ACA	0.30 (3888)	0.54 (9255)	p<0.0001
	GUC*	0.23 (2373)	0.05 (675)	p<0.0001		ACG	0.01 (75)	0.02 (279)	p<0.0001
	GUA	0.18 (1931)	0.34 (4415)	p<0.0001	**Tyr**	UAU	0.61 (3454)	0.89 (12575)	p<0.0001
	GUG	0.02 (240)	0.05 (604)	p<0.0001		UAC*	0.39 (2185)	0.11 (1602)	p<0.0001
Pro	CCU	0.07 (709)	0.17 (1772)	p<0.0001	**Ala**	GCU*	0.43 (4808)	0.34 (2362)	p<0.0001
	CCC	0.01 (160)	0.03 (305)	p<0.0001		GCC*	0.27 (2957)	0.08 (586)	p<0.0001
	CCA*	0.91 (9942)	0.79 (8450)	p<0.0001		GCA	0.30 (3347)	0.56 (3971)	p<0.0001
	CCG	0.01 (62)	0.01 (152)	p<0.0001		GCG	0.00 (40)	0.02 (112)	p<0.0001
His	CAU	0.64 (2448)	0.90 (4589)	p<0.0001	**Ile**	AUU	0.57 (6463)	0.63 (21177)	p<0.0001
	CAC*	0.36 (1389)	0.10 (524)	p<0.0001		AUC*	0.33 (3773)	0.08 (2749)	p<0.0001
Gln	CAA*	0.98 (12416)	0.97 (10902)	p<0.0001		AUA	0.10 (1079)	0.29 (9629)	p<0.0001
	CAG	0.02 (278)	0.03 (387)	p<0.0001	**Asp**	GAU	0.85 (7245)	0.93 (15479)	p<0.0001
Asn	AAU	0.72 (9549)	0.93 (37190)	p<0.0001		GAC*	0.15 (1322)	0.07 (1199)	p<0.0001
	AAC*	0.28 (3720)	0.07 (2789)	p<0.0001	**Glu**	GAA	0.87 (9889)	0.90 (15860)	p<0.0001
Lys	AAA	0.79 (11482)	0.90 (23502)	p<0.0001		GAG*	0.13 (1424)	0.10 (1769)	p<0.0001
	AAG*	0.21 (3102)	0.10 (2729)	p<0.0001	**Gly**	GGU*	0.83 (8681)	0.68 (9687)	p<0.0001
Cys	UGU	0.81 (1691)	0.90 (5130)	p<0.0001		GGC*	0.05 (499)	0.04 (552)	p=0.0007
	UGC*	0.19 (384)	0.10 (550)	p<0.0001		GGA	0.11 (1170)	0.25 (3546)	p<0.0001
Arg	CGU*	0.41 (2763)	0.15 (1089)	p<0.0001		GGG	0.01 (68)	0.03 (371)	p<0.0001
	CGC	0.00 (8)	0.00 (32)	p=0.0003					
	CGA	0.00 (34)	0.02 (178)	p<0.0001					
	CGG	0.00 (4)	0.00 (18)	p=0.0042					
	AGA	0.57 (3912)	0.75 (5400)	p<0.0001					
	AGG	0.01 (93)	0.06 (461)	p<0.0001					

Table 45 | Codon Usage from Expression Data for the Amoebozoan *D. discoideum*

Amino Acids	Codons	Top 5% Expressed Genes	Bottom 5% Expressed Genes	Chi-Sqaured
Phe	UUU	0.32 (5378)	0.37 (4708)	p<0.0001
	UUC*	0.68 (11271)	0.63 (7847)	p<0.0001
Leu	UUA	0.10 (4002)	0.12 (3353)	p<0.0001
	UUG	0.21 (8290)	0.20 (5886)	p=0.0011
	CUU	0.09 (3319)	0.11 (3194)	p<0.0001
	CUC*	0.28 (11016)	0.25 (7201)	p<0.0001
	CUA	0.08 (2993)	0.10 (2825)	p<0.0001
	CUG*	0.24 (9266)	0.23 (6555)	p=0.0002
Val	GUU	0.15 (4098)	0.17 (3303)	p<0.0001
	GUC*	0.40 (10795)	0.37 (7208)	p<0.0001
	GUA	0.14 (3857)	0.17 (3352)	p<0.0001
	GUG	0.31 (8282)	0.29 (5780)	p=0.0048
Pro	CCU	0.25 (4969)	0.28 (4408)	p<0.0001
	CCC*	0.34 (6929)	0.27 (4313)	p<0.0001
	CCA	0.27 (5399)	0.30 (4871)	p<0.0001
	CCG	0.15 (2980)	0.15 (2423)	p=0.2498
His	CAU	0.34 (3457)	0.39 (3315)	p<0.0001
	CAC*	0.66 (6724)	0.61 (5153)	p<0.0001
Gln	CAA	0.47 (8262)	0.49 (7060)	p<0.0001
	CAG*	0.53 (9470)	0.51 (7350)	p<0.0001
Asn	AAU	0.37 (7348)	0.40 (6240)	p<0.0001
	AAC*	0.63 (12576)	0.60 (9347)	p<0.0001
Lys	AAA	0.32 (7794)	0.38 (7048)	p<0.0001
	AAG*	0.68 (16194)	0.62 (11440)	p<0.0001
Cys	UGU	0.48 (3292)	0.50 (2923)	p=0.0826
	UGC	0.52 (3546)	0.50 (2960)	p=0.0826
Arg	CGU	0.11 (2589)	0.12 (2171)	p=0.0277
	CGC	0.10 (2436)	0.11 (2110)	p=0.0010
	CGA	0.17 (3943)	0.17 (3118)	p=0.9409
	CGG	0.05 (1132)	0.06 (1068)	p<0.0001
	AGA*	0.38 (8859)	0.36 (6583)	p<0.0001
	AGG	0.19 (4332)	0.18 (3338)	p=0.2430
Ser	UCU	0.22 (8549)	0.21 (6252)	p=0.0132
	UCC*	0.19 (7489)	0.18 (5181)	p<0.0001
	UCA	0.16 (6053)	0.18 (5323)	p<0.0001
	UCG	0.14 (5437)	0.14 (4039)	p=0.2810
	AGU	0.12 (4674)	0.13 (3741)	p=0.0088
	AGC	0.16 (6247)	0.16 (4626)	p=0.1778
Thr	ACU	0.19 (5463)	0.21 (4784)	p<0.0001
	ACC*	0.30 (8358)	0.28 (6233)	p<0.0001
	ACA	0.26 (7337)	0.28 (6221)	p<0.0001
	ACG*	0.25 (7113)	0.23 (5032)	p<0.0001
Tyr	UAU	0.35 (4519)	0.40 (4213)	p<0.0001
	UAC*	0.65 (8305)	0.60 (6219)	p<0.0001
Ala	GCU	0.23 (6228)	0.26 (5380)	p<0.0001
	GCC*	0.37 (9848)	0.31 (6374)	p<0.0001
	GCA	0.25 (6794)	0.27 (5538)	p<0.0001
	GCG	0.15 (3942)	0.15 (3043)	p=0.4278
Ile	AUU	0.18 (4411)	0.21 (3859)	p<0.0001
	AUC*	0.68 (16978)	0.61 (11181)	p<0.0001
	AUA	0.14 (3453)	0.18 (3229)	p<0.0001
Asp	GAU	0.52 (11492)	0.51 (8688)	p=0.0318
	GAC	0.48 (10761)	0.49 (8498)	p=0.0318
Glu	GAA	0.43 (11720)	0.46 (8964)	p<0.0001
	GAG*	0.57 (15820)	0.54 (10620)	p<0.0001
Gly	GGU	0.28 (7152)	0.27 (4780)	p=0.0952
	GGC	0.19 (4956)	0.21 (3708)	p<0.0001
	GGA*	0.43 (11093)	0.40 (7055)	p<0.0001
	GGG	0.10 (2616)	0.12 (2176)	p<0.0001

Table 46 | Codon Usage from Expression Data for the Amoebozoan *P. fungivorum*

Amino Acids	Codons	Top 5% Expressed Genes	Bottom 5% Expressed Genes	Chi-Sqaured	Amino Acids	Codons	Top 5% Expressed Genes	Bottom 5% Expressed Genes	Chi-Sqaured
Phe	UUU	0.47 (2892)	0.60 (5147)	p<0.0001	**Ser**	UCU	0.08 (1161)	0.08 (2382)	p=0.5244
	UUC*	0.53 (3297)	0.40 (3496)	p<0.0001		UCC*	0.16 (2178)	0.13 (3721)	p<0.0001
Leu	UUA	0.00 (27)	0.00 (126)	p=0.0028		UCA	0.06 (899)	0.09 (2544)	p<0.0001
	UUG	0.04 (626)	0.06 (2281)	p<0.0001		UCG*	0.58 (8072)	0.45 (12760)	p<0.0001
	CUU	0.14 (1992)	0.16 (5572)	p<0.0001		AGU	0.02 (306)	0.05 (1405)	p<0.0001
	CUC*	0.54 (7748)	0.44 (15911)	p<0.0001		AGC	0.10 (1377)	0.19 (5278)	p<0.0001
	CUA	0.01 (175)	0.03 (1147)	p<0.0001	**Thr**	ACU	0.11 (1187)	0.15 (2772)	p<0.0001
	CUG	0.26 (3778)	0.30 (10878)	p<0.0001		ACC*	0.50 (5204)	0.43 (8127)	p<0.0001
Val	GUU	0.13 (1700)	0.15 (4107)	p<0.0001		ACA	0.07 (745)	0.13 (2491)	p<0.0001
	GUC*	0.51 (6861)	0.47 (12868)	p<0.0001		ACG*	0.31 (3262)	0.29 (5425)	p<0.0001
	GUA	0.02 (220)	0.04 (1163)	p<0.0001	**Tyr**	UAU	0.09 (406)	0.21 (1261)	p<0.0001
	GUG	0.35 (4671)	0.34 (9378)	p=0.1990		UAC*	0.91 (3965)	0.79 (4767)	p<0.0001
Pro	CCU	0.11 (974)	0.12 (2305)	p=0.0003	**Ala**	GCU	0.17 (3248)	0.17 (9947)	p=0.0927
	CCC*	0.23 (2085)	0.19 (3518)	p<0.0001		GCC*	0.46 (9075)	0.40 (23127)	p<0.0001
	CCA	0.08 (693)	0.15 (2858)	p<0.0001		GCA	0.10 (1866)	0.14 (8182)	p<0.0001
	CCG*	0.58 (5270)	0.54 (10063)	p<0.0001		GCG	0.28 (5414)	0.29 (16948)	p<0.0001
His	CAU	0.23 (795)	0.32 (2371)	p<0.0001	**Ile**	AUU	0.24 (1806)	0.26 (2752)	p=0.0120
	CAC*	0.77 (2711)	0.68 (5124)	p<0.0001		AUC*	0.75 (5596)	0.71 (7602)	p<0.0001
Gln	CAA	0.09 (477)	0.18 (1655)	p<0.0001		AUA	0.01 (68)	0.03 (304)	p<0.0001
	CAG*	0.91 (5130)	0.82 (7555)	p<0.0001	**Asp**	GAU	0.21 (2147)	0.31 (6049)	p<0.0001
Asn	AAU	0.08 (470)	0.18 (1264)	p<0.0001		GAC*	0.79 (8240)	0.69 (13231)	p<0.0001
	AAC*	0.92 (5424)	0.82 (5872)	p<0.0001	**Glu**	GAA	0.08 (769)	0.15 (2835)	p<0.0001
Lys	AAA	0.03 (269)	0.10 (849)	p<0.0001		GAG*	0.92 (8960)	0.85 (16070)	p<0.0001
	AAG*	0.97 (8443)	0.90 (7362)	p<0.0001	**Gly**	GGU	0.16 (2169)	0.15 (3523)	p=0.0048
Cys	UGU	0.15 (421)	0.23 (889)	p<0.0001		GGC*	0.66 (8753)	0.60 (13787)	p<0.0001
	UGC*	0.85 (2387)	0.77 (3055)	p<0.0001		GGA	0.07 (955)	0.10 (2373)	p<0.0001
Arg	CGU*	0.14 (1295)	0.12 (2722)	p<0.0001		GGG	0.10 (1360)	0.15 (3392)	p<0.0001
	CGC*	0.52 (4801)	0.43 (10029)	p<0.0001					
	CGA	0.04 (383)	0.10 (2251)	p<0.0001					
	CGG	0.25 (2343)	0.29 (6846)	p<0.0001					
	AGA	0.02 (150)	0.02 (513)	p=0.0008					
	AGG	0.03 (325)	0.04 (1008)	p=0.0008					

Table 47 | Codon Usage from Expression Data for the Apusomonadan *T. trahens*

Amino Acids	Codons	Top 5% Expressed Genes	Bottom 5% Expressed Genes	Chi-Sqaured
Phe	UUU	0.09 (318)	0.19 (1252)	p<0.0001
	UUC*	0.91 (3337)	0.81 (5448)	p<0.0001
Leu	UUA*	0.02 (120)	0.00 (76)	p<0.0001
	UUG*	0.03 (236)	0.02 (2135)	p<0.0001
	CUU*	0.09 (647)	0.01 (1631)	p<0.0001
	CUC*	0.42 (3180)	0.05 (5807)	p<0.0001
	CUA	0.01 (75)	0.00 (323)	p=0.0061
	CUG	0.44 (3302)	0.91 (13037)	p<0.0001
Val	GUU*	0.14 (934)	0.07 (957)	p<0.0001
	GUC*	0.63 (4104)	0.38 (5106)	p<0.0001
	GUA	0.01 (70)	0.02 (212)	p=0.0061
	GUG	0.21 (1356)	0.53 (7200)	p<0.0001
Pro	CCU*	0.10 (490)	0.06 (1436)	p<0.0001
	CCC*	0.54 (2735)	0.35 (8602)	p<0.0001
	CCA	0.03 (173)	0.07 (1634)	p<0.0001
	CCG	0.32 (1629)	0.52 (12719)	p<0.0001
His	CAU	0.21 (449)	0.32 (2387)	p<0.0001
	CAC*	0.79 (1644)	0.68 (5073)	p<0.0001
Gln	CAA	0.04 (127)	0.17 (1553)	p<0.0001
	CAG*	0.96 (2753)	0.83 (7732)	p<0.0001
Asn	AAU	0.12 (414)	0.39 (1679)	p<0.0001
	AAC*	0.88 (3158)	0.61 (2623)	p<0.0001
Lys	AAA	0.04 (198)	0.09 (355)	p<0.0001
	AAG*	0.96 (4627)	0.91 (3487)	p<0.0001
Cys	UGU	0.17 (244)	0.18 (1175)	p=0.2460
	UGC	0.83 (1188)	0.82 (5230)	p=0.2460
Arg	CGU*	0.19 (1032)	0.07 (1295)	p<0.0001
	CGC*	0.64 (3528)	0.41 (7301)	p<0.0001
	CGA	0.02 (133)	0.07 (1190)	p<0.0001
	CGG	0.10 (577)	0.41 (7381)	p<0.0001
	AGA*	0.02 (94)	0.01 (172)	p<0.0001
	AGG	0.03 (145)	0.03 (626)	p=0.0019
Ser	UCU*	0.14 (939)	0.04 (873)	p<0.0001
	UCC*	0.50 (3258)	0.29 (6097)	p<0.0001
	UCA	0.02 (98)	0.03 (702)	p<0.0001
	UCG	0.15 (997)	0.31 (6399)	p<0.0001
	AGU	0.02 (131)	0.05 (1088)	p<0.0001
	AGC	0.17 (1080)	0.27 (5520)	p<0.0001
Thr	ACU*	0.13 (651)	0.07 (844)	p<0.0001
	ACC*	0.79 (4090)	0.62 (7805)	p<0.0001
	ACA	0.02 (126)	0.07 (886)	p<0.0001
	ACG	0.06 (307)	0.24 (3070)	p<0.0001
Tyr	UAU	0.15 (369)	0.81 (1198)	p<0.0001
	UAC*	0.85 (2093)	0.19 (277)	p<0.0001
Ala	GCU*	0.17 (1550)	0.08 (2520)	p<0.0001
	GCC*	0.72 (6789)	0.65 (21443)	p<0.0001
	GCA	0.02 (213)	0.06 (1973)	p<0.0001
	GCG	0.09 (828)	0.22 (7179)	p<0.0001
Ile	AUU	0.20 (877)	0.20 (1310)	p=0.2534
	AUC	0.79 (3524)	0.77 (4931)	p=0.0457
	AUA	0.02 (78)	0.02 (158)	p=0.0104
Asp	GAU	0.30 (1455)	0.32 (4037)	p=0.0428
	GAC	0.70 (3351)	0.68 (8631)	p=0.0428
Glu	GAA	0.04 (193)	0.14 (1402)	p<0.0001
	GAG*	0.96 (4697)	0.86 (8942)	p<0.0001
Gly	GGU*	0.33 (2276)	0.08 (1938)	p<0.0001
	GGC	0.55 (3747)	0.55 (13808)	p=0.5406
	GGA	0.03 (201)	0.05 (1371)	p<0.0001
	GGG	0.09 (588)	0.32 (8177)	p<0.0001

Table 48 | Codon Usage from Expression Data for the Opisthokont *F. alba*

Amino Acids	Codons	Top 5% Expressed Genes	Bottom 5% Expressed Genes	Chi-Sqaured
Phe	UUU	0.37 (2411)	0.49 (5304)	p<0.0001
	UUC*	0.63 (4063)	0.51 (5517)	p<0.0001
Leu	UUA	0.00 (63)	0.02 (612)	p<0.0001
	UUG	0.11 (1480)	0.21 (5889)	p<0.0001
	CUU	0.15 (2004)	0.17 (4901)	p<0.0001
	CUC*	0.47 (6364)	0.27 (7830)	p<0.0001
	CUA	0.01 (123)	0.04 (1042)	p<0.0001
	CUG	0.25 (3393)	0.29 (8434)	p<0.0001
Val	GUU	0.21 (2373)	0.27 (5852)	p<0.0001
	GUC*	0.58 (6394)	0.36 (7937)	p<0.0001
	GUA	0.01 (157)	0.05 (1034)	p<0.0001
	GUG	0.19 (2115)	0.32 (7087)	p<0.0001
Pro	CCU	0.15 (1228)	0.24 (4019)	p<0.0001
	CCC*	0.50 (3988)	0.26 (4435)	p<0.0001
	CCA	0.12 (931)	0.23 (3967)	p<0.0001
	CCG	0.23 (1812)	0.27 (4675)	p<0.0001
His	CAU	0.17 (613)	0.38 (2585)	p<0.0001
	CAC*	0.83 (2952)	0.62 (4293)	p<0.0001
Gln	CAA	0.30 (1970)	0.45 (6055)	p<0.0001
	CAG*	0.70 (4550)	0.55 (7416)	p<0.0001
Asn	AAU	0.13 (815)	0.33 (3922)	p<0.0001
	AAC*	0.87 (5617)	0.67 (7834)	p<0.0001
Lys	AAA	0.07 (636)	0.31 (3210)	p<0.0001
	AAG*	0.93 (7847)	0.69 (7293)	p<0.0001
Cys	UGU	0.10 (265)	0.24 (1168)	p<0.0001
	UGC*	0.90 (2328)	0.76 (3753)	p<0.0001
Arg	CGU	0.19 (1476)	0.18 (2973)	p=0.2181
	CGC*	0.63 (4972)	0.41 (6714)	p<0.0001
	CGA	0.10 (777)	0.20 (3265)	p<0.0001
	CGG	0.05 (366)	0.12 (1983)	p<0.0001
	AGA	0.02 (135)	0.05 (749)	p<0.0001
	AGG	0.02 (187)	0.05 (830)	p<0.0001
Ser	UCU	0.15 (1829)	0.15 (4452)	p=0.1192
	UCC*	0.25 (2941)	0.19 (5917)	p<0.0001
	UCA	0.05 (581)	0.10 (3067)	p<0.0001
	UCG*	0.32 (3861)	0.25 (7716)	p<0.0001
	AGU	0.03 (391)	0.08 (2398)	p<0.0001
	AGC	0.20 (2383)	0.22 (6814)	p<0.0001
Thr	ACU	0.23 (2186)	0.22 (4463)	p=0.1281
	ACC*	0.53 (5009)	0.34 (6816)	p<0.0001
	ACA	0.07 (653)	0.18 (3617)	p<0.0001
	ACG	0.18 (1686)	0.26 (5262)	p<0.0001
Tyr	UAU	0.16 (707)	0.30 (1991)	p<0.0001
	UAC*	0.84 (3742)	0.70 (4582)	p<0.0001
Ala	GCU*	0.26 (4332)	0.24 (8670)	p=0.0007
	GCC*	0.49 (8326)	0.31 (10950)	p<0.0001
	GCA	0.11 (1777)	0.21 (7603)	p<0.0001
	GCG	0.14 (2446)	0.24 (8460)	p<0.0001
Ile	AUU	0.41 (3143)	0.50 (6448)	p<0.0001
	AUC*	0.58 (4386)	0.46 (5885)	p<0.0001
	AUA	0.01 (58)	0.04 (472)	p<0.0001
Asp	GAU	0.24 (2030)	0.39 (6221)	p<0.0001
	GAC*	0.76 (6257)	0.61 (9622)	p<0.0001
Glu	GAA	0.21 (1740)	0.39 (5778)	p<0.0001
	GAG*	0.79 (6455)	0.61 (9130)	p<0.0001
Gly	GGU	0.20 (2071)	0.22 (3737)	p<0.0001
	GGC*	0.71 (7407)	0.49 (8234)	p<0.0001
	GGA	0.07 (713)	0.18 (3087)	p<0.0001
	GGG	0.03 (310)	0.11 (1819)	p<0.0001

Table 49 | Codon Usage from Expression Data for the Opisthokont *C. owczarzaki*

Amino Acids	Codons	Top 5% Expressed Genes	Bottom 5% Expressed Genes	Chi-Sqaured	Amino Acids	Codons	Top 5% Expressed Genes	Bottom 5% Expressed Genes	Chi-Sqaured
Phe	UUU	0.28 (2704)	0.48 (3290)	p<0.0001	**Ser**	UCU*	0.19 (3224)	0.16 (2851)	p<0.0001
	UUC*	0.72 (6815)	0.52 (3516)	p<0.0001		UCC*	0.18 (3086)	0.13 (2406)	p<0.0001
Leu	UUA	0.07 (1377)	0.11 (1969)	p<0.0001		UCA	0.16 (2725)	0.17 (3134)	p=0.0017
	UUG	0.21 (4103)	0.21 (3778)	p=0.0787		UCG	0.15 (2510)	0.16 (2947)	p=0.0001
	CUU*	0.17 (3377)	0.14 (2505)	p<0.0001		AGU	0.17 (2938)	0.20 (3627)	p<0.0001
	CUC	0.15 (2929)	0.16 (2901)	p<0.0001		AGC	0.15 (2532)	0.18 (3187)	p<0.0001
	CUA	0.14 (2828)	0.13 (2374)	p=0.0267	**Thr**	ACU*	0.25 (3879)	0.23 (4023)	p<0.0001
	CUG*	0.26 (5244)	0.23 (4125)	p<0.0001		ACC*	0.34 (5256)	0.24 (4195)	p<0.0001
Val	GUU	0.22 (4178)	0.22 (3231)	p=0.4428		ACA	0.24 (3734)	0.31 (5412)	p<0.0001
	GUC*	0.23 (4321)	0.21 (3097)	p<0.0001		ACG	0.16 (2521)	0.22 (3738)	p<0.0001
	GUA	0.21 (3963)	0.23 (3383)	p<0.0001	**Tyr**	UAU	0.30 (2357)	0.44 (2943)	p<0.0001
	GUG	0.34 (6523)	0.35 (5206)	p=0.2985		UAC*	0.70 (5460)	0.56 (3709)	p<0.0001
Pro	CCU*	0.28 (3079)	0.24 (2683)	p<0.0001	**Ala**	GCU*	0.29 (6666)	0.24 (4190)	p<0.0001
	CCC*	0.38 (4199)	0.24 (2605)	p<0.0001		GCC*	0.34 (7656)	0.26 (4546)	p<0.0001
	CCA	0.20 (2151)	0.30 (3265)	p<0.0001		GCA	0.19 (4332)	0.28 (4946)	p<0.0001
	CCG	0.14 (1482)	0.23 (2496)	p<0.0001		GCG	0.18 (4169)	0.23 (4028)	p<0.0001
His	CAU	0.30 (1640)	0.43 (2641)	p<0.0001	**Ile**	AUU	0.35 (4643)	0.35 (3734)	p=0.4880
	CAC*	0.70 (3845)	0.57 (3473)	p<0.0001		AUC*	0.54 (7134)	0.35 (3705)	p<0.0001
Gln	CAA	0.35 (3169)	0.46 (4170)	p<0.0001		AUA	0.11 (1497)	0.29 (3106)	p<0.0001
	CAG*	0.65 (5764)	0.54 (4961)	p<0.0001	**Asp**	GAU	0.53 (8172)	0.53 (7422)	p=0.6754
Asn	AAU	0.37 (3915)	0.49 (4931)	p<0.0001		GAC	0.47 (7289)	0.47 (6685)	p=0.6754
	AAC*	0.63 (6779)	0.51 (5229)	p<0.0001	**Glu**	GAA	0.33 (5501)	0.46 (5779)	p<0.0001
Lys	AAA	0.26 (4285)	0.42 (4588)	p<0.0001		GAG*	0.67 (11309)	0.54 (6728)	p<0.0001
	AAG*	0.74 (12082)	0.58 (6223)	p<0.0001	**Gly**	GGU*	0.37 (7114)	0.31 (4398)	p<0.0001
Cys	UGU	0.48 (2113)	0.52 (2583)	p=0.0005		GGC	0.26 (4976)	0.27 (3813)	p=0.0628
	UGC*	0.52 (2258)	0.48 (2391)	p=0.0005		GGA*	0.29 (5545)	0.23 (3365)	p<0.0001
Arg	CGU*	0.31 (3584)	0.18 (2299)	p<0.0001		GGG	0.09 (1718)	0.19 (2752)	p<0.0001
	CGC*	0.24 (2771)	0.19 (2384)	p<0.0001					
	CGA	0.16 (1840)	0.16 (2043)	p=0.4326					
	CGG	0.06 (693)	0.13 (1704)	p<0.0001					
	AGA	0.16 (1835)	0.19 (2457)	p<0.0001					
	AGG	0.08 (931)	0.14 (1757)	p<0.0001					

Table 50 | Codon Usage from Expression Data for the Opisthokont *S. arctica*

Amino Acids	Codons	Top 5% Expressed Genes	Bottom 5% Expressed Genes	Chi-Sqaured	Amino Acids	Codons	Top 5% Expressed Genes	Bottom 5% Expressed Genes	Chi-Sqaured
Phe	UUU	0.41 (2315)	0.57 (4227)	p<0.0001	**Ser**	UCU	0.23 (3043)	0.25 (5146)	p<0.0001
	UUC*	0.59 (3346)	0.43 (7412)	p<0.0001		UCC*	0.20 (2638)	0.16 (3262)	p<0.0001
Leu	UUA	0.05 (770)	0.11 (2542)	p<0.0001		UCA	0.10 (1396)	0.16 (3263)	p<0.0001
	UUG	0.11 (1761)	0.14 (3318)	p<0.0001		UCG*	0.14 (1823)	0.10 (2106)	p<0.0001
	CUU*	0.26 (4137)	0.24 (5720)	p<0.0001		AGU	0.11 (1529)	0.14 (2807)	p<0.0001
	CUC*	0.30 (4779)	0.19 (4475)	p<0.0001		AGC*	0.22 (3000)	0.19 (3899)	p<0.0001
	CUA	0.08 (1285)	0.14 (3257)	p<0.0001	**Thr**	ACU	0.23 (2412)	0.27 (3652)	p<0.0001
	CUG*	0.20 (3193)	0.18 (4377)	p<0.0001		ACC*	0.26 (2779)	0.23 (3166)	p<0.0001
Val	GUU	0.27 (3061)	0.29 (3617)	p=0.0394		ACA	0.26 (2794)	0.32 (4446)	p<0.0001
	GUC*	0.38 (4229)	0.26 (3316)	p<0.0001		ACG*	0.25 (2641)	0.18 (2445)	p<0.0001
	GUA	0.11 (1210)	0.19 (2430)	p<0.0001	**Tyr**	UAU	0.42 (2425)	0.53 (3740)	p<0.0001
	GUG	0.24 (2656)	0.26 (3262)	p=0.0003		UAC*	0.58 (3344)	0.47 (3345)	p<0.0001
Pro	CCU	0.28 (2141)	0.28 (2880)	p=0.3067	**Ala**	GCU	0.25 (3817)	0.26 (4508)	p<0.0001
	CCC	0.25 (1974)	0.24 (2435)	p=0.0204		GCC*	0.30 (4601)	0.24 (4170)	p<0.0001
	CCA	0.23 (1820)	0.31 (3148)	p<0.0001		GCA	0.26 (4027)	0.34 (5783)	p<0.0001
	CCG*	0.24 (1831)	0.17 (1724)	p<0.0001		GCG*	0.20 (3097)	0.15 (2562)	p<0.0001
His	CAU	0.37 (1411)	0.46 (2636)	p<0.0001	**Ile**	AUU	0.32 (3042)	0.36 (4439)	p<0.0001
	CAC*	0.63 (2387)	0.54 (3042)	p<0.0001		AUC*	0.46 (4443)	0.33 (4061)	p<0.0001
Gln	CAA	0.28 (1802)	0.42 (3762)	p<0.0001		AUA	0.22(2081)	0.31 (3959)	p<0.0001
	CAG*	0.72 (4604)	0.58 (5160)	p<0.0001	**Asp**	GAU	0.43 (4499)	0.53 (5945)	p<0.0001
Asn	AAU	0.39 (2966)	0.51 (4443)	p<0.0001		GAC*	0.57 (6070)	0.47 (5267)	p<0.0001
	AAC*	0.61 (4548)	0.49 (4313)	p<0.0001	**Glu**	GAA	0.28 (3366)	0.42 (5202)	p<0.0001
Lys	AAA	0.20 (2084)	0.35 (3615)	p<0.0001		GAG*	0.72 (8523)	0.58 (7251)	p<0.0001
	AAG*	0.80 (8122)	0.65 (6802)	p<0.0001	**Gly**	GGU	0.18 (2109)	0.20 (2085)	p=0.0115
Cys	UGU	0.36 (1647)	0.42 (2124)	p<0.0001		GGC*	0.35 (3972)	0.27 (2883)	p<0.0001
	UGC*	0.64 (2992)	0.58 (2944)	p<0.0001		GGA	0.24 (2752)	0.30 (3124)	p<0.0001
Arg	CGU*	0.18 (1638)	0.14 (1776)	p<0.0001		GGG	0.23 (2635)	0.23 (2472)	p=0.4667
	CGC*	0.32 (2976)	0.19 (2309)	p<0.0001					
	CGA	0.08 (749)	0.14 (1662)	p<0.0001					
	CGG	0.09 (871)	0.15 (1828)	p<0.0001					
	AGA	0.14 (1279)	0.21 (2561)	p<0.0001					
	AGG*	0.19 (1815)	0.18 (2175)	p=0.0008					

Table 51 | Codon Usage from Expression Data for the Excavate *G. lamblia*

Amino Acids	Codons	Top 5% Expressed Genes	Bottom 5% Expressed Genes	Chi-Sqaured
Phe	UUU	0.36 (1892)	0.78 (5488)	p<0.0001
	UUC*	0.64 (3408)	0.22 (1527)	p<0.0001
Leu	UUA	0.18 (1755)	0.46 (5896)	p<0.0001
	UUG	0.10 (952)	0.13 (1647)	p<0.0001
	CUU*	0.25 (2419)	0.17 (2212)	p<0.0001
	CUC*	0.30 (2866)	0.04 (559)	p<0.0001
	CUA	0.05 (457)	0.14 (1871)	p<0.0001
	CUG*	0.13 (1223)	0.06 (772)	p<0.0001
Val	GUU	0.40 (3346)	0.42 (2508)	p=0.0891
	GUC*	0.26 (2131)	0.09 (568)	p<0.0001
	GUA	0.19 (1544)	0.36 (2164)	p<0.0001
	GUG*	0.16 (1304)	0.13 (786)	p<0.0001
Pro	CCU	0.37 (1836)	0.38 (1185)	p=0.9037
	CCC*	0.30 (1474)	0.09 (276)	p<0.0001
	CCA	0.23 (1141)	0.44 (1400)	p<0.0001
	CCG	0.09 (455)	0.09 (292)	p=0.9839
His	CAU	0.42 (970)	0.77 (1516)	p<0.0001
	CAC*	0.58 (1334)	0.23 (449)	p<0.0001
Glu	GAA	0.55 (4347)	0.74 (4623)	p<0.0001
	GAG*	0.45 (3539)	0.26 (1600)	p<0.0001
Asn	AAU	0.44 (2546)	0.77 (7201)	p<0.0001
	AAC*	0.56 (3252)	0.23 (2125)	p<0.0001
Lys	AAA	0.40 (3490)	0.75 (7281)	p<0.0001
	AAG*	0.60 (5327)	0.25 (2475)	p<0.0001
Cys	UGU	0.30 (1229)	0.51 (2046)	p<0.0001
	UGC*	0.70 (2820)	0.49 (1988)	p<0.0001
Arg	CGU*	0.29 (1570)	0.11 (401)	p<0.0001
	CGC*	0.22 (1164)	0.08 (281)	p<0.0001
	CGA	0.02 (95)	0.13 (469)	p<0.0001
	CGG	0.04 (218)	0.06 (198)	p=0.0007
	AGA	0.32 (1705)	0.47 (1658)	p<0.0001
	AGG	0.11 (596)	0.14 (492)	p<0.0001
Ser	UCU*	0.31 (2445)	0.27 (2600)	p<0.0001
	UCC*	0.23 (1832)	0.07 (685)	p<0.0001
	UCA	0.15 (1163)	0.28 (2683)	p<0.0001
	UCG*	0.09 (679)	0.06 (604)	p<0.0001
	AGU	0.07 (556)	0.21 (1974)	p<0.0001
	AGC*	0.15 (1217)	0.10 (968)	p<0.0001
Thr	ACU	0.37 (2580)	0.37 (2442)	p=0.9621
	ACC*	0.35 (2439)	0.14 (955)	p<0.0001
	ACA	0.15 (1057)	0.39 (2560)	p<0.0001
	ACG*	0.13 (940)	0.10 (674)	p<0.0001
Tyr	UAU	0.53 (1995)	0.78 (4485)	p<0.0001
	UAC*	0.47 (1743)	0.22 (1278)	p<0.0001
Ala	GCU	0.38 (3917)	0.36 (1598)	p=0.0082
	GCC*	0.31 (3176)	0.12 (534)	p<0.0001
	GCA	0.17 (1708)	0.40 (1789)	p<0.0001
	GCG*	0.14 (1399)	0.11 (508)	p=0.0002
Ile	AUU	0.41 (3236)	0.49 (6198)	p<0.0001
	AUC*	0.46 (3694)	0.11 (1428)	p<0.0001
	AUA	0.13 (1034)	0.40 (5040)	p<0.0001
Asp	GAU	0.52 (3454)	0.78 (4660)	p<0.0001
	GAC*	0.48 (3162)	0.22 (1311)	p<0.0001
Gln	CAA	0.15 (972)	0.14 (1760)	p=0.0806
	CAG*	0.42 (2718)	0.08 (942)	p<0.0001
	UUA	0.25 (1619)	0.58 (7179)	p<0.0001
	UAG	0.17 (1088)	0.20 (2484)	p<0.0001
Gly	GGU*	0.44 (3626)	0.34 (1355)	p<0.0001
	GGC*	0.31 (2557)	0.15 (594)	p<0.0001
	GGA	0.17 (1419)	0.39 (1548)	p<0.0001
	GGG	0.09 (710)	0.12 (497)	p<0.0001

Table 52 | Codon Usage from Expression Data for the Excavate *S. salmonicida*

Amino Acids	Codons	Top 5% Expressed Genes	Bottom 5% Expressed Genes	Chi-Sqaured
Phe	UUU	0.22 (944)	0.41 (3441)	p<0.0001
	UUC*	0.78 (3302)	0.59 (5031)	p<0.0001
Leu	UUA	0.01 (69)	0.02 (685)	p<0.0001
	UUG	0.06 (573)	0.13 (3839)	p<0.0001
	CUU	0.11 (970)	0.13 (3888)	p<0.0001
	CUC	0.26 (2366)	0.27 (8182)	p=0.0861
	CUA	0.03 (268)	0.05 (1608)	p<0.0001
	CUG*	0.54 (4916)	0.41 (12412)	p<0.0001
Val	GUU	0.10 (892)	0.12 (2906)	p<0.0001
	GUC	0.24 (2054)	0.26 (6244)	p=0.0001
	GUA	0.04 (321)	0.08 (1848)	p<0.0001
	GUG*	0.62 (5298)	0.54 (12935)	p<0.0001
Pro	CCU	0.12 (660)	0.14 (3678)	p=0.0003
	CCC*	0.26 (1401)	0.18 (4743)	p<0.0001
	CCA	0.10 (529)	0.19 (4879)	p<0.0001
	CCG*	0.52 (2823)	0.49 (12900)	p<0.0001
His	CAU	0.16 (464)	0.26 (2692)	p<0.0001
	CAC*	0.84 (2476)	0.74 (7607)	p<0.0001
Gln	CAA	0.09 (370)	0.20 (2649)	p<0.0001
	CAG*	0.91 (3744)	0.80 (10339)	p<0.0001
Asn	AAU	0.11 (438)	0.21 (1503)	p<0.0001
	AAC*	0.89 (3641)	0.79 (5614)	p<0.0001
Lys	AAA	0.05 (389)	0.18 (1352)	p<0.0001
	AAG*	0.95 (7004)	0.82 (6007)	p<0.0001
Cys	UGU	0.16 (383)	0.26 (1860)	p<0.0001
	UGC*	0.84 (1986)	0.74 (5371)	p<0.0001
Arg	CGU	0.15 (1181)	0.15 (3796)	p=0.9635
	CGC*	0.65 (5209)	0.41 (10512)	p<0.0001
	CGA	0.03 (267)	0.11 (2751)	p<0.0001
	CGG	0.09 (702)	0.21 (5291)	p<0.0001
	AGA	0.02 (197)	0.04 (1143)	p<0.0001
	AGG	0.06 (444)	0.09 (2257)	p<0.0001
Ser	UCU	0.14 (1078)	0.13 (6188)	p=0.0029
	UCC*	0.25 (1991)	0.20 (9773)	p<0.0001
	UCA	0.04 (336)	0.07 (3281)	p<0.0001
	UCG	0.27 (2079)	0.25 (12398)	p=0.0087
	AGU	0.04 (319)	0.06 (2886)	p<0.0001
	AGC	0.26 (2033)	0.30 (14781)	p<0.0001
Thr	ACU	0.10 (620)	0.13 (2686)	p<0.0001
	ACC*	0.31 (2013)	0.28 (6038)	p<0.0001
	ACA	0.11 (722)	0.20 (4164)	p<0.0001
	ACG*	0.48 (3092)	0.39 (8371)	p<0.0001
Tyr	UAU	0.10 (356)	0.23 (1552)	p<0.0001
	UAC*	0.90 (3189)	0.77 (5265)	p<0.0001
Ala	GCU	0.15 (1654)	0.14 (7000)	p=0.0013
	GCC*	0.35 (3830)	0.26 (13106)	p<0.0001
	GCA	0.08 (894)	0.17 (8394)	p<0.0001
	GCG	0.41 (4509)	0.43 (21493)	p=0.0026
Ile	AUU	0.20 (1011)	0.28 (2091)	p<0.0001
	AUC*	0.75 (3777)	0.60 (4531)	p<0.0001
	AUA	0.04 (218)	0.13 (975)	p<0.0001
Asp	GAU	0.24 (1358)	0.30 (4792)	p<0.0001
	GAC*	0.76 (4197)	0.70 (11092)	p<0.0001
Glu	GAA	0.10 (691)	0.21 (3566)	p<0.0001
	GAG*	0.90 (6503)	0.79 (13620)	p<0.0001
Gly	GGU*	0.22 (1769)	0.20 (4478)	p<0.0001
	GGC*	0.66 (5304)	0.51 (11638)	p<0.0001
	GGA	0.05 (372)	0.11 (2505)	p<0.0001
	GGG	0.08 (652)	0.19 (4311)	p<0.0001

Table 53 | Codon Usage from Expression Data for the Excavate *L. major*

Amino Acids	Codons	Top 5% Expressed Genes	Bottom 5% Expressed Genes	Chi-Sqaured
Phe	UUU	0.16 (1099)	0.41 (4510)	p<0.0001
	UUC*	0.84 (5822)	0.59 (6558)	p<0.0001
Leu	UUA	0.01 (133)	0.03 (963)	p<0.0001
	UUG	0.05 (760)	0.13 (4143)	p<0.0001
	CUU	0.08 (1087)	0.11 (3438)	p<0.0001
	CUC	0.21 (2907)	0.27 (8467)	p<0.0001
	CUA	0.02 (213)	0.04 (1325)	p<0.0001
	CUG*	0.63 (8753)	0.42 (13133)	p<0.0001
Val	GUU	0.09 (1159)	0.12 (2927)	p<0.0001
	GUC	0.22 (2943)	0.27 (6859)	p<0.0001
	GUA	0.02 (324)	0.07 (1675)	p<0.0001
	GUG*	0.67 (8898)	0.54 (13654)	p<0.0001
Pro	CCU	0.11 (874)	0.15 (3393)	p<0.0001
	CCC*	0.29 (2359)	0.25 (5526)	p<0.0001
	CCA	0.05 (427)	0.16 (3545)	p<0.0001
	CCG*	0.55 (4457)	0.44 (9719)	p<0.0001
His	CAU	0.08 (333)	0.20 (1878)	p<0.0001
	CAC*	0.92 (3867)	0.80 (7473)	p<0.0001
Gln	CAA	0.09 (567)	0.23 (3258)	p<0.0001
	CAG*	0.91 (5542)	0.77 (10618)	p<0.0001
Asn	AAU	0.07 (460)	0.21 (2158)	p<0.0001
	AAC*	0.93 (6281)	0.79 (8094)	p<0.0001
Lys	AAA	0.04 (541)	0.24 (2614)	p<0.0001
	AAG*	0.96 (12137)	0.76 (8451)	p<0.0001
Cys	UGU	0.15 (569)	0.26 (1582)	p<0.0001
	UGC*	0.85 (3149)	0.74 (4620)	p<0.0001
Arg	CGU	0.14 (1734)	0.18 (4203)	p<0.0001
	CGC*	0.73 (8747)	0.42 (10171)	p<0.0001
	CGA	0.03 (319)	0.11 (2547)	p<0.0001
	CGG	0.06 (683)	0.20 (4685)	p<0.0001
	AGA	0.02 (186)	0.04 (866)	p<0.0001
	AGG	0.03 (336)	0.06 (1484)	p<0.0001
Ser	UCU	0.10 (1086)	0.11 (3551)	p<0.0001
	UCC*	0.28 (3213)	0.23 (7239)	p<0.0001
	UCA	0.02 (270)	0.08 (2363)	p<0.0001
	UCG*	0.29 (3295)	0.24 (7542)	p<0.0001
	AGU	0.03 (374)	0.08 (2552)	p<0.0001
	AGC	0.27 (3082)	0.26 (8049)	p=0.0018
Thr	ACU	0.07 (764)	0.09 (1944)	p<0.0001
	ACC	0.32 (3288)	0.33 (7255)	p=0.0012
	ACA	0.06 (622)	0.14 (3091)	p<0.0001
	ACG*	0.55 (5732)	0.43 (9426)	p<0.0001
Tyr	UAU	0.07 (371)	0.15 (1229)	p<0.0001
	UAC*	0.93 (4702)	0.85 (7058)	p<0.0001
Ala	GCU*	0.13 (2124)	0.11 (4782)	p<0.0001
	GCC*	0.38 (6239)	0.29 (12262)	p<0.0001
	GCA	0.05 (855)	0.23 (9717)	p<0.0001
	GCG*	0.44 (7389)	0.37 (15624)	p<0.0001
Ile	AUU	0.17 (1402)	0.32 (3272)	p<0.0001
	AUC*	0.80 (6537)	0.60 (6160)	p<0.0001
	AUA	0.03 (215)	0.07 (764)	p<0.0001
Asp	GAU	0.22 (1921)	0.31 (5054)	p<0.0001
	GAC*	0.78 (6808)	0.69 (11339)	p<0.0001
Glu	GAA	0.08 (818)	0.24 (4551)	p<0.0001
	GAG*	0.92 (9899)	0.76 (14283)	p<0.0001
Gly	GGU*	0.22 (2840)	0.21 (4389)	p=0.0002
	GGC*	0.67 (8510)	0.49 (10376)	p<0.0001
	GGA	0.04 (510)	0.11 (2348)	p<0.0001
	GGG	0.06 (819)	0.19 (4104)	p<0.0001

Table 54 | Codon Usage from Expression Data for the Excavate *L. pyrrhocoris*

Amino Acids	Codons	Top 5% Expressed Genes	Bottom 5% Expressed Genes	Chi-Sqaured
Phe	UUU*	0.44 (3759)	0.31 (1136)	p<0.0001
	UUC	0.56 (4750)	0.69 (2580)	p<0.0001
Leu	UUA	0.04 (937)	0.03 (246)	p<0.0001
	UUG*	0.15 (3814)	0.12 (1089)	p<0.0001
	CUU	0.13 (3156)	0.12 (1148)	p=0.6980
	CUC	0.24 (5877)	0.23 (2130)	p=0.4658
	CUA	0.06 (1382)	0.05 (428)	p=0.0013
	CUG	0.39 (9828)	0.45 (4165)	p<0.0001
Val	GUU	0.14 (2624)	0.13 (999)	p=0.0779
	GUC	0.23 (4288)	0.23 (1733)	p=0.9629
	GUA*	0.10 (1811)	0.07 (527)	p<0.0001
	GUG	0.54 (10228)	0.57 (4409)	p<0.0001
Pro	CCU*	0.18 (2854)	0.15 (849)	p<0.0001
	CCC	0.24 (3822)	0.28 (1592)	p<0.0001
	CCA*	0.20 (3197)	0.15 (871)	p<0.0001
	CCG	0.38 (6124)	0.42 (2381)	p<0.0001
His	CAU*	0.28 (1985)	0.23 (625)	p<0.0001
	CAC	0.72 (5184)	0.77 (2041)	p<0.0001
Gln	CAA*	0.23 (2573)	0.18 (658)	p<0.0001
	CAG	0.77 (8785)	0.82 (2988)	p<0.0001
Asn	AAU*	0.28 (2238)	0.20 (636)	p<0.0001
	AAC	0.72 (5824)	0.80 (2518)	p<0.0001
Lys	AAA*	0.21 (1998)	0.16 (632)	p<0.0001
	AAG	0.79 (7416)	0.84 (3341)	p<0.0001
Cys	UGU	0.26 (1230)	0.23 (533)	p=0.0064
	UGC	0.74 (3589)	0.77 (1828)	p=0.0064
Arg	CGU*	0.19 (3608)	0.16 (1157)	p<0.0001
	CGC	0.41 (7828)	0.45 (3297)	p<0.0001
	CGA*	0.11 (2098)	0.08 (568)	p<0.0001
	CGG	0.18 (3510)	0.17 (1254)	p=0.0172
	AGA	0.04 (796)	0.05 (363)	p=0.0051
	AGG	0.07 (1378)	0.10 (732)	p<0.0001
Ser	UCU*	0.14 (3576)	0.11 (943)	p<0.0001
	UCC	0.19 (4706)	0.21 (1743)	p<0.0001
	UCA*	0.10 (2384)	0.08 (633)	p<0.0001
	UCG	0.21 (5131)	0.23 (1943)	p<0.0001
	AGU*	0.10 (2409)	0.08 (634)	p<0.0001
	AGC	0.27 (6667)	0.29 (2374)	p=0.0008
Thr	ACU*	0.13 (2120)	0.11 (665)	p<0.0001
	ACC	0.30 (4948)	0.29 (1763)	p=0.0591
	ACA*	0.18 (2998)	0.15 (899)	p<0.0001
	ACG	0.39 (6309)	0.45 (2769)	p<0.0001
Tyr	UAU*	0.21 (1444)	0.18 (478)	p<0.0001
	UAC	0.79 (5315)	0.82 (2243)	p<0.0001
Ala	GCU	0.17 (5177)	0.16 (1734)	p=0.0086
	GCC	0.28 (8473)	0.30 (3224)	p=0.0005
	GCA*	0.20 (6055)	0.15 (1658)	p<0.0001
	GCG	0.35 (10390)	0.39 (4157)	p<0.0001
Ile	AUU*	0.35 (2962)	0.26 (1010)	p<0.0001
	AUC	0.56 (4707)	0.67 (2569)	p<0.0001
	AUA*	0.09 (795)	0.07 (280)	p<0.0001
Asp	GAU*	0.38 (4925)	0.30 (1347)	p<0.0001
	GAC	0.62 (8154)	0.70 (3129)	p<0.0001
Glu	GAA*	0.24 (4048)	0.19 (1045)	p<0.0001
	GAG	0.76 (12672)	0.81 (4416)	p<0.0001
Gly	GGU	0.21 (3626)	0.21 (1444)	p=0.7543
	GGC	0.48 (8174)	0.52 (3584)	p<0.0001
	GGA*	0.12 (2066)	0.09 (651)	p<0.0001
	GGG	0.19 (3254)	0.17 (1198)	p=0.0043

Table 55 | Codon Usage from Expression Data for the Excavate *L. seymouri*

Amino Acids	Codons	Top 5% Expressed Genes	Bottom 5% Expressed Genes	Chi-Sqaured
Phe	UUU	0.55 (2053)	0.57 (6540)	p=0.1093
	UUC	0.45 (1674)	0.43 (5015)	p=0.1093
Leu	UUA	0.10 (686)	0.10 (3382)	p=0.5767
	UUG	0.20 (1363)	0.21 (7248)	p=0.0463
	CUU	0.24 (1642)	0.25 (8373)	p=0.6590
	CUC*	0.19 (1246)	0.15 (5110)	p<0.0001
	CUA	0.07 (482)	0.09 (3151)	p<0.0001
	CUG	0.19 (1291)	0.19 (6588)	p=0.6880
Val	GUU	0.30 (1749)	0.31 (8325)	p=0.0124
	GUC	0.16 (945)	0.15 (4045)	p=0.1284
	GUA	0.15 (855)	0.17 (4433)	p<0.0001
	GUG*	0.40 (2347)	0.37 (9760)	p<0.0001
Pro	CCU	0.23 (812)	0.26 (4201)	p<0.0001
	CCC*	0.26 (943)	0.13 (2058)	p<0.0001
	CCA	0.27 (976)	0.33 (5173)	p<0.0001
	CCG	0.23 (832)	0.28 (4435)	p<0.0001
His	CAU	0.41 (784)	0.48 (4320)	p<0.0001
	CAC*	0.59 (1112)	0.52 (4700)	p<0.0001
Gln	CAA	0.44 (1194)	0.46 (6171)	p=0.0284
	CAG	0.56 (1549)	0.54 (7297)	p=0.0296
Asn	AAU	0.44 (1359)	0.51 (6570)	p<0.0001
	AAC*	0.56 (1695)	0.49 (6384)	p<0.0001
Lys	AAA	0.34 (1821)	0.44 (6587)	p<0.0001
	AAG*	0.66 (3586)	0.56 (8271)	p<0.0001
Cys	UGU	0.47 (770)	0.51 (3898)	p=0.0120
	UGC	0.53 (858)	0.49 (3781)	p=0.0120
Arg	CGU	0.25 (1403)	0.24 (5955)	p=0.0194
	CGC*	0.32 (1804)	0.20 (4944)	p<0.0001
	CGA	0.10 (541)	0.14 (3447)	p<0.0001
	CGG	0.13 (717)	0.18 (4430)	p<0.0001
	AGA	0.08 (468)	0.10 (2625)	p<0.0001
	AGG	0.12 (649)	0.15 (3781)	p<0.0001
Ser	UCU	0.18 (958)	0.16 (4948)	p=0.0016
	UCC*	0.17 (927)	0.15 (4730)	p=0.0005
	UCA	0.16 (857)	0.17 (5245)	p=0.0473
	UCG	0.14 (765)	0.15 (4534)	p=0.3406
	AGU	0.16 (837)	0.19 (5899)	p<0.0001
	AGC	0.19 (1007)	0.17 (5263)	p=0.0040
Thr	ACU	0.22 (1013)	0.24 (4814)	p=0.0691
	ACC	0.20 (913)	0.22 (4502)	p=0.0042
	ACA	0.29 (1300)	0.29 (5955)	p=0.4763
	ACG*	0.29 (1304)	0.25 (5093)	p<0.0001
Tyr	UAU	0.44 (1007)	0.45 (3577)	p=0.1826
	UAC	0.56 (1301)	0.55 (4332)	p=0.1826
Ala	GCU	0.26 (1658)	0.27 (7905)	p=0.1216
	GCC*	0.25 (1565)	0.21 (6027)	p<0.0001
	GCA	0.25 (1598)	0.28 (8262)	p<0.0001
	GCG	0.24 (1510)	0.24 (7021)	p=0.6681
Ile	AUU	0.48 (1742)	0.46 (5949	p=0.0255
	AUC*	0.33 (1210)	0.28 (3625)	p<0.0001
	AUA	0.19 (685)	0.26 (3418)	p<0.0001
Asp	GAU	0.51 (1856)	0.57 (10284)	p<0.0001
	GAC*	0.49 (1750)	0.43 (7650)	p<0.0001
Glu	GAA	0.44 (2051)	0.46 (11368)	p=0.0371
	GAG	0.56 (2578)	054 (13354)	p=0.0371
Gly	GGU*	0.37 (1957)	0.33 (7798)	p<0.0001
	GGC*	0.24 (1265)	0.22 (5096)	p=0.0003
	GGA	0.22 (1178)	0.23 (5349)	p=0.5105
	GGG	0.17 (890)	0.23 (5322)	p<0.0001

Table 56 | Codon Usage from Expression Data for the Excavate *T. brucei*

Amino Acids	Codons	Top 5% Expressed Genes	Bottom 5% Expressed Genes	Chi-Sqaured
Phe	UUU	0.61 (8601)	0.60 (9097)	p=0.0424
	UUC	0.39 (5572)	0.40 (6187)	p=0.0424
Leu	UUA	0.05 (2000)	0.06 (2358)	p=0.1344
	UUG	0.19 (6786)	0.20 (8223)	p<0.0001
	CUU*	0.22 (7880)	0.18 (7238)	p<0.0001
	CUC*	0.19 (6832)	0.16 (6504)	p<0.0001
	CUA	0.04 (1492)	0.04 (1702)	p=0.7502
	CUG	0.32 (11496)	0.37 (15113)	p<0.0001
Val	GUU*	0.24 (7534)	0.20 (8032)	p<0.0001
	GUC*	0.18 (5790)	0.16 (6574)	p<0.0001
	GUA*	0.09 (2802)	0.07 (2696)	p<0.0001
	GUG	0.49 (15292)	0.57 (23113)	p<0.0001
Pro	CCU*	0.20 (4554)	0.18 (4412)	p<0.0001
	CCC	0.23 (5171)	0.23 (5467)	p=0.2859
	CCA*	0.28 (6303)	0.24 (5772)	p<0.0001
	CCG	0.29 (6532)	0.36 (8643)	p<0.0001
His	CAU	0.37 (3713)	0.37 (4584)	p=0.6860
	CAC	0.63 (6321)	0.63 (7895)	p=0.6860
Gln	CAA*	0.38 (5844)	0.35 (4897)	p<0.0001
	CAG	0.62 (9625)	0.65 (9007)	p<0.0001
Asn	AAU*	0.46 (7634)	0.41 (6106)	p<0.0001
	AAC	0.54 (8833)	0.59 (8894)	p<0.0001
Lys	AAA	0.35 (8442)	0.44 (5356)	p<0.0001
	AAG*	0.65 (15532)	0.56 (6752)	p<0.0001
Cys	UGU*	0.38 (3532)	0.32 (4340)	p<0.0001
	UGC	0.62 (5654)	0.68 (9230)	p<0.0001
Arg	CGU*	0.22 (5775)	0.18 (5880)	p<0.0001
	CGC*	0.29 (7769)	0.25 (8083)	p<0.0001
	CGA	0.09 (2456)	0.12 (3845)	p<0.0001
	CGG	0.14 (3722)	0.20 (6524)	p<0.0001
	AGA	0.10 (2523)	0.09 (2855)	p=0.0024
	AGG	0.16 (4236)	0.16 (5257)	p=0.5039

Amino Acids	Codons	Top 5% Expressed Genes	Bottom 5% Expressed Genes	Chi-Sqaured
Ser	UCU*	0.16 (5428)	0.13 (5370)	p<0.0001
	UCC*	0.20 (6455)	0.18 (7557)	p<0.0001
	UCA*	0.15 (4897)	0.12 (4797)	p<0.0001
	UCG	0.15 (5050)	0.22 (9360)	p<0.0001
	AGU*	0.15 (4818)	0.12 (4898)	p<0.0001
	AGC	0.19 (6365)	0.23 (9721)	p<0.0001
Thr	ACU*	0.18 (5328)	0.16 (4447)	p<0.0001
	ACC*	0.22 (6387)	0.19 (5503)	p<0.0001
	ACA*	0.27 (7776)	0.21 (5889)	p<0.0001
	ACG	0.34 (9846)	0.45 (12752)	p<0.0001
Tyr	UAU	0.31 (2916)	0.31 (3009)	p=0.8264
	UAC	0.69 (6533)	0.69 (6692)	p=0.8264
Ala	GCU*	0.21 (7954)	0.19 (7954)	p<0.0001
	GCC*	0.26 (10109)	0.22 (9538)	p<0.0001
	GCA*	0.24 (9075)	0.20 (8341)	p<0.0001
	GCG	0.30 (11466)	0.39 (16767)	p<0.0001
Ile	AUU*	0.48 (7385)	0.45 (6733)	p<0.0001
	AUC	0.38 (5913)	0.39 (5737)	p=0.7943
	AUA	0.14 (2089)	0.16 (2400)	p<0.0001
Asp	GAU*	0.47 (9699)	0.40 (8189)	p<0.0001
	GAC	0.53 (10965)	0.60 (12068)	p<0.0001
Glu	GAA*	0.41 (12162)	0.37 (7987)	p<0.0001
	GAG	0.59 (17604)	0.63 (13455)	p<0.0001
Gly	GGU*	0.26 (8445)	0.21 (8052)	p<0.0001
	GGC	0.34 (11194)	0.40 (15315)	p<0.0001
	GGA*	0.22 (7204)	0.17 (6352)	p<0.0001
	GGG	0.19 (6231)	0.22 (8181)	p<0.0001

Table 57 | Codon Usage from Expression Data for the Excavate *T. cruzi*

Amino Acids	Codons	Top 5% Expressed Genes	Bottom 5% Expressed Genes	Chi-Sqaured
Phe	UUU	0.46 (2162)	0.75 (6947)	p<0.0001
	UUC*	0.54 (2551)	0.25 (2297)	p<0.0001
Leu	UUA	0.07 (737)	0.26 (5526)	p<0.0001
	UUG	0.15 (1669)	0.17 (3615)	p=0.0003
	CUU	0.21 (2276)	0.28 (5978)	p<0.0001
	CUC*	0.24 (2643)	0.11 (2453)	p<0.0001
	CUA	0.04 (469)	0.09 (1861)	p<0.0001
	CUG*	0.29 (3224)	0.10 (2184)	p<0.0001
Val	GUU	0.24 (2632)	0.35 (5333)	p<0.0001
	GUC*	0.15 (1660)	0.09 (1431)	p<0.0001
	GUA	0.11 (1240)	0.27 (4122)	p<0.0001
	GUG*	0.49 (5254)	0.29 (4491)	p<0.0001
Pro	CCU	0.23 (1493)	0.35 (4144)	p<0.0001
	CCC*	0.27 (1754)	0.15 (1804)	p<0.0001
	CCA	0.32 (2064)	0.36 (4325)	p<0.0001
	CCG*	0.18 (1145)	0.13 (1583)	p<0.0001
His	CAU	0.37 (1238)	0.66 (3741)	p<0.0001
	CAC*	0.63 (2083)	0.34 (1891)	p<0.0001
Gln	CAA	0.37 (2033)	0.65 (5231)	p<0.0001
	CAG*	0.63 (3504)	0.35 (2859)	p<0.0001
Asn	AAU	0.45 (2881)	0.73 (8074)	p<0.0001
	AAC*	0.55 (3535)	0.27 (2957)	p<0.0001
Lys	AAA	0.28 (3028)	0.65 (6749)	p<0.0001
	AAG*	0.72 (7896)	0.35 (3686)	p<0.0001
Cys	UGU	0.44 (1251)	0.71 (3489)	p<0.0001
	UGC*	0.56 (1577)	0.29 (1415)	p<0.0001
Arg	CGU	0.29 (2507)	0.31 (4179)	p<0.0001
	CGC*	0.39 (3380)	0.12 (1540)	p<0.0001
	CGA	0.08 (690)	0.19 (2575)	p<0.0001
	CGG	0.08 (659)	0.10 (1388)	p<0.0001
	AGA	0.09 (787)	0.19 (2546)	p<0.0001
	AGG	0.09 (754)	0.09 (1138)	p=0.8614
Ser	UCU	0.20 (1836)	0.31 (6168)	p<0.0001
	UCC*	0.18 (1697)	0.12 (2289)	p<0.0001
	UCA	0.15 (1353)	0.21 (4105)	p<0.0001
	UCG*	0.10 (895)	0.06 (1116)	p<0.0001
	AGU	0.18 (1595)	0.24 (4760)	p<0.0001
	AGC*	0.20 (1832)	0.07 (1415)	p<0.0001
Thr	ACU	0.22 (1939)	0.31 (4307)	p<0.0001
	ACC*	0.23 (2033)	0.15 (2135)	p<0.0001
	ACA	0.34 (3020)	0.40 (5565)	p<0.0001
	ACG*	0.21 (1885)	0.15 (2071)	p<0.0001
Tyr	UAU	0.33 (1257)	0.67 (4148)	p<0.0001
	UAC*	0.67 (2496)	0.33 (2061)	p<0.0001
Ala	GCU	0.26 (3362)	0.32 (4541)	p<0.0001
	GCC*	0.29 (3795)	0.19 (2770)	p<0.0001
	GCA	0.25 (3310)	0.33 (4682)	p<0.0001
	GCG*	0.20 (2540)	0.17 (2372)	p<0.0001
Ile	AUU	0.49 (2982)	0.56 (6634)	p<0.0001
	AUC*	0.40 (2465)	0.15 (1778)	p<0.0001
	AUA	0.11 (669)	0.29 (3364)	p<0.0001
Asp	GAU	0.53 (3712)	0.80 (8052)	p<0.0001
	GAC*	0.47 (3301)	0.20 (2022)	p<0.0001
Glu	GAA	0.41 (4508)	0.65 (9015)	p<0.0001
	GAG*	0.59 (6518)	0.35 (4793)	p<0.0001
Gly	GGU	0.39 (3684)	0.41 (5630)	p=0.0240
	GGC*	0.28 (2649)	0.12 (1631)	p<0.0001
	GGA	0.22 (2023)	0.33 (4497)	p<0.0001
	GGG	0.11 (1022)	0.15 (2050)	p<0.0001

Table 58 | Codon Usage from Expression Data for the Excavate *T. theileri*

Amino Acids	Codons	Top 5% Expressed Genes	Bottom 5% Expressed Genes	Chi-Sqaured
Phe	UUU	0.39 (7731)	0.75 (24652)	p<0.0001
	UUC*	0.61 (11944)	0.25 (24652)	p<0.0001
Leu	UUA	0.36 (12918)	0.43 (22934)	p<0.0001
	UUG*	0.18 (6546)	0.15 (8105)	p<0.0001
	CUU*	0.27 (9518)	0.20 (10635)	p<0.0001
	CUC*	0.15 (5267)	0.06 (2994)	p<0.0001
	CUA	0.04 (1287)	0.12 (6549)	p<0.0001
	CUG	0.01 (272)	0.03 (1746)	p<0.0001
Val	GUU*	0.59 (14152)	0.50 (12866)	p<0.0001
	GUC*	0.24 (5841)	0.11 (2774)	p<0.0001
	GUA	0.15 (3551)	0.31 (7905)	p<0.0001
	GUG	0.02 (535)	0.09 (2351)	p<0.0001
Pro	CCU*	0.62 (8978)	0.55 (9173)	p<0.0001
	CCC*	0.23 (3367)	0.07 (1202)	p<0.0001
	CCA	0.14 (2039)	0.36 (6082)	p<0.0001
	CCG	0.00 (50)	0.02 (325)	p<0.0001
His	CAU	0.49 (3843)	0.73 (4736)	p<0.0001
	CAC*	0.51 (4069)	0.27 (1778)	p<0.0001
Glu	GAA*	0.92 (26061)	0.80 (21646)	p<0.0001
	GAG	0.08 (2393)	0.20 (5388)	p<0.0001
Asn	AAU	0.56 (15200)	0.76 (41366)	p<0.0001
	AAC*	0.44 (11987)	0.24 (13032)	p<0.0001
Lys	AAA	0.46 (16579)	0.77 (33687)	p<0.0001
	AAG*	0.54 (19405)	0.23 (10330)	p<0.0001
Cys	UGU	0.47 (2946)	0.58 (11706)	p<0.0001
	UGC*	0.53 (3349)	0.42 (8557)	p<0.0001
Arg	CGU*	0.15 (2346)	0.07 (915)	p<0.0001
	CGC	0.02 (299)	0.03 (403)	p<0.0001
	CGA	0.00 (49)	0.03 (417)	p<0.0001
	CGG	0.00 (10)	0.00 (51)	p<0.0001
	AGA*	0.80 (12642)	0.73 (9473)	p<0.0001
	AGG	0.02 (380)	0.13 (1709)	p<0.0001
Ser	UCU*	0.40 (10861)	0.33 (17787)	p<0.0001
	UCC*	0.12 (3221)	0.05 (2628)	p<0.0001
	UCA	0.19 (5107)	0.26 (14099)	p<0.0001
	UCG	0.01 (214)	0.02 (1283)	p<0.0001
	AGU	0.15 (4233)	0.21 (11329)	p<0.0001
	AGC	0.13 (3676)	0.13 (7261)	p=0.6670
Thr	ACU*	0.59 (12722)	0.52 (16455)	p<0.0001
	ACC*	0.23 (4908)	0.10 (3036)	p<0.0001
	ACA	0.18 (3817)	0.36 (11412)	p<0.0001
	ACG	0.00 (76)	0.02 (772)	p<0.0001
Tyr	UAU	0.53 (8548)	0.72 (20880)	p<0.0001
	UAC*	0.47 (7505)	0.28 (8250)	p<0.0001
Ala	GCU*	0.67 (17095)	0.56 (11986)	p<0.0001
	GCC*	0.17 (4383)	0.09 (1963)	p<0.0001
	GCA	0.15 (3931)	0.33 (7000)	p<0.0001
	GCG	0.00 (91)	0.03 (540)	p<0.0001
Ile	AUU*	0.58 (16730)	0.51 (26894)	p<0.0001
	AUC*	0.29 (8425)	0.14 (7185)	p<0.0001
	AUA	0.13 (3765)	0.36 (19042)	p<0.0001
Asp	GAU	0.75 (16719)	0.81 (23196)	p<0.0001
	GAC*	0.25 (5658)	0.19 (5316)	p<0.0001
Gln	CAA*	0.35 (8619)	0.57 (11214)	p<0.0001
	CAG	0.02 (379)	0.04 (2425)	p<0.0001
	UAA	0.52 (12966)	0.60 (33372)	p<0.0001
	UAG	0.12 (2886)	0.15 (8623)	p<0.0001
Gly	GGU*	0.66 (15266)	0.37 (8998)	p<0.0001
	GGC	0.08 (1946)	0.11 (2729)	p<0.0001
	GGA	0.25 (5689)	0.46 (11272)	p<0.0001
	GGG	0.01 (250)	0.06 (1597)	p<0.0001

Table 59 | Codon Usage from Expression Data for the Alveolate *T. thermophila*

Amino Acids	Codons	Top 5% Expressed Genes	Bottom 5% Expressed Genes	Chi-Sqaured
Phe	UUU	0.62 (12588)	0.80 (28737)	p<0.0001
	UUC*	0.38 (7623)	0.20 (7401)	p<0.0001
Leu	UUA	0.30 (12420)	0.40 (24208)	p<0.0001
	UUG*	0.15 (6097)	0.14 (8444)	p<0.0001
	CUU*	0.34 (13960)	0.26 (15751)	p<0.0001
	CUC*	0.07 (2976)	0.05 (2994)	p<0.0001
	CUA	0.10 (3999)	0.12 (7285)	p<0.0001
	CUG*	0.04 (1590)	0.03 (1935)	p<0.0001
Val	GUU	0.48 (15344)	0.49 (13726)	p=0.1963
	GUC*	0.16 (5206)	0.10 (2792)	p<0.0001
	GUA	0.22 (7128)	0.29 (8097)	p<0.0001
	GUG	0.13 (4101)	0.12 (3505)	p=0.1063
Pro	CCU	0.48 (9491)	0.48 (8908)	p=0.4135
	CCC	0.04 (749)	0.05 (898)	p<0.0001
	CCA*	0.44 (8696)	0.42 (7896)	p=0.0004
	CCG	0.04 (733)	0.05 (920)	p<0.0001
His	CAU	0.79 (7216)	0.83 (8977)	p<0.0001
	CAC*	0.21 (1935)	0.17 (1875)	p<0.0001
Gln	CAA	0.75 (12646)	0.79 (15352)	p<0.0001
	CAG*	0.25 (4247)	0.21 (4108)	p<0.0001
Asn	AAU	0.67 (16248)	0.76 (35011)	p<0.0001
	AAC*	0.33 (8060)	0.24 (11225)	p<0.0001
Lys	AAA	0.73 (28206)	0.82 (48351)	p<0.0001
	AAG*	0.27 (10541)	0.18 (10359)	p<0.0001
Cys	UGU*	0.66 (5637)	0.56 (6978)	p<0.0001
	UGC	0.34 (2896)	0.44 (5428)	p<0.0001
Arg	CGU*	0.15 (3405)	0.11 (2200)	p<0.0001
	CGC	0.03 (775)	0.03 (626)	p=0.0520
	CGA	0.06 (1364)	0.09 (1790)	p<0.0001
	CGG	0.01 (323)	0.02 (461)	p<0.0001
	AGA*	0.57 (12893)	0.55 (11046)	p<0.0001
	AGG	0.17 (3792)	0.20 (4062)	p<0.0001
Ser	UCU	0.24 (7992)	0.28 (13913)	p<0.0001
	UCC	0.05 (1711)	0.06 (2828)	p=0.0001
	UCA*	0.32 (10588)	0.24 (11742)	p<0.0001
	UCG	0.04 (1341)	0.04 (2112)	p=0.0495
	AGU	0.24 (7937)	0.26 (13030)	p<0.0001
	AGC	0.12 (3942)	0.12 (5727)	p=0.4844
Thr	ACU*	0.49 (13441)	0.44 (13637)	p<0.0001
	ACC*	0.12 (3391)	0.10 (3017)	p<0.0001
	ACA	0.36 (9842)	0.43 (13204)	p<0.0001
	ACG	0.03 (872)	0.04 (1127)	p=0.0017
Tyr	UAU	0.74 (12653)	0.79 (23953)	p<0.0001
	UAC*	0.26 (4561)	0.21 (6186)	p<0.0001
Ala	GCU*	0.56 (18538)	0.50 (11159)	p<0.0001
	GCC*	0.11 (3529)	0.08 (1894)	p<0.0001
	GCA	0.28 (9355)	0.36 (7989)	p<0.0001
	GCG	0.04 (1420)	0.06 (1350)	p<0.0001
Ile	AUU*	0.51 (16118)	0.45 (28363)	p<0.0001
	AUC*	0.23 (7072)	0.14 (9053)	p<0.0001
	AUA	0.26 (8221)	0.41 (25948)	p<0.0001
Asp	GAU	0.76 (21854)	0.81 (23854)	p<0.0001
	GAC*	0.24 (6959)	0.19 (5524)	p<0.0001
Glu	GAA	0.81 (30170)	0.81 (32998)	p=0.7174
	GAG	0.19 (7141)	0.19 (7759)	p=0.7174
Gly	GGU*	0.33 (10705)	0.28 (6587)	p<0.0001
	GGC*	0.12 (3835)	0.10 (2383)	p<0.0001
	GGA	0.40 (12832)	0.41 (9773)	p=0.0004
	GGG	0.16 (5030)	0.21 (5043)	p<0.0001

Table 60 | Codon Usage from Expression Data for the Alveolate *S. coeruleus*

Amino Acids	Codons	Top 5% Expressed Genes	Bottom 5% Expressed Genes	Chi-Sqaured
Phe	UUU	0.88 (4437)	0.87 (4697)	p=0.2110
	UUC	0.12 (625)	0.13 (712)	p=0.2110
Leu	UUA*	0.70 (6938)	0.56 (6530)	p<0.0001
	UUG	0.12 (1217)	0.16 (1887)	p<0.0001
	CUU	0.10 (989)	0.13 (1547)	p<0.0001
	CUC	0.01 (95)	0.02 (282)	p<0.0001
	CUA	0.07 (658)	0.10 (1176)	p<0.0001
	CUG	0.01 (85)	0.02 (186)	p<0.0001
Val	GUU*	0.49 (2884)	0.44 (2041)	p<0.0001
	GUC	0.04 (233)	0.03 (157)	p=0.1353
	GUA	0.40 (2383)	0.43 (1993)	p=0.0052
	GUG	0.07 (411)	0.10 (442)	p<0.0001
Pro	CCU	0.38 (1262)	0.48 (2140)	p<0.0001
	CCC	0.05 (165)	0.07 (298)	p=0.0011
	CCA*	0.56 (1849)	0.42 (1841)	p<0.0001
	CCG	0.01 (48)	0.03 (142)	p<0.0001
His	CAU	0.85 (1865)	0.86 (2301)	p=0.2691
	CAC	0.15 (336)	0.14 (379)	p=0.2691
Gln	CAA	0.88 (3039)	0.87 (3080)	p=0.0446
	CAG	0.12 (409)	0.13 (479)	p=0.0446
Asn	AAU	0.87 (11249)	0.87 (12607)	p=0.0855
	AAC	0.13 (1748)	0.13 (1842)	p=0.0855
Lys	AAA*	0.86 (12090)	0.83 (14090)	p<0.0001
	AAG	0.14 (1937)	0.17 (2894)	p<0.0001
Cys	UGU	0.79 (1441)	0.75 (1818)	p=0.0078
	UGC	0.21 (390)	0.25 (599)	p=0.0078
Arg	CGU	0.08 (370)	0.10 (437)	p=0.0057
	CGC	0.01 (47)	0.01 (67)	p=0.0463
	CGA	0.04 (197)	0.05 (243)	p=0.0138
	CGG	0.00 (9)	0.01 (42)	p<0.0001
	AGA*	0.77 (3592)	0.67 (3057)	p<0.0001
	AGG	0.09 (427)	0.15 (704)	p<0.0001
Ser	UCU	0.26 (2122)	0.33 (3797)	p<0.0001
	UCC	0.05 (407)	0.07 (825)	p<0.0001
	UCA*	0.32 (2656)	0.28 (3304)	p<0.0001
	UCG	0.02 (188)	0.03 (338)	p=0.0060
	AGU*	0.29 (2404)	0.24 (2829)	p<0.0001
	AGC*	0.06 (507)	0.05 (554)	p<0.0001
Thr	ACU	0.38 (2078)	0.37 (3179)	p=0.2227
	ACC	0.06 (349)	0.07 (636)	p=0.0215
	ACA	0.51 (2840)	0.50 (4365)	p=0.1809
	ACG	0.05 (266)	0.06 (519)	p=0.0032
Tyr	UAU*	0.87 (4403)	0.84 (4779)	p<0.0001
	UAC	0.13 (672)	0.16 (908)	p<0.0001
Ala	GCU*	0.42 (1979)	0.36 (1119)	p<0.0001
	GCC	0.06 (282)	0.06 (187)	p=0.9961
	GCA	0.50 (2342)	0.52 (1638)	p=0.0194
	GCG	0.02 (112)	0.06 (184)	p<0.0001
Ile	AUU*	0.49 (4811)	0.43 (5447)	p<0.0001
	AUC	0.05 (450)	0.06 (751)	p<0.0001
	AUA	0.46 (4479)	0.52 (6600)	p<0.0001
Asp	GAU*	0.90 (7023)	0.86 (6294)	p<0.0001
	GAC	0.10 (790)	0.14 (1005)	p<0.0001
Glu	GAA	0.86 (6941)	0.85 (12895)	p=0.0910
	GAG	0.14 (1156)	0.15 (2294)	p=0.0910
Gly	GGU*	0.38 (2056)	0.29 (1081)	p<0.0001
	GGC	0.03 (141)	0.05 (197)	p<0.0001
	GGA	0.55 (3003)	0.59 (2209)	p=0.0003
	GGG	0.05 (279)	0.08 (285)	p<0.0001

Table 61 | Codon Usage from Expression Data for the Alveolate *P. gallinaceum*

Amino Acids	Codons	Top 5% Expressed Genes	Bottom 5% Expressed Genes	Chi-Sqaured
Phe	UUU	0.78 (2868)	0.83 (2904)	p<0.0001
	UUC*	0.22 (792)	0.17 (588)	p<0.0001
Leu	UUA*	0.66 (4729)	0.55 (4392)	p<0.0001
	UUG	0.13 (915)	0.17 (1372)	p<0.0001
	CUU	0.12 (880)	0.13 (1024)	p=0.2167
	CUC	0.03 (184)	0.03 (244)	p=0.0559
	CUA	0.06 (395)	0.09 (676)	p<0.0001
	CUG	0.01 (77)	0.03 (215)	p<0.0001
Val	GUU	0.43 (2165)	0.42 (1920)	p=0.2268
	GUC	0.07 (343)	0.08 (363)	p=0.0409
	GUA*	0.43 (2146)	0.36 (1671)	p<0.0001
	GUG	0.07 (374)	0.14 (635)	p<0.0001
Pro	CCU	0.29 (696)	0.41 (955)	p<0.0001
	CCC	0.09 (205)	0.10 (233)	p=0.1273
	CCA*	0.59 (1411)	0.45 (1060)	p<0.0001
	CCG	0.03 (71)	0.05 (109)	p=0.0030
His	CAU	0.76 (1409)	0.83 (1519)	p<0.0001
	CAC*	0.24 (440)	0.17 (317)	p<0.0001
Gln	CAA	0.89 (2409)	0.87 (2653)	p=0.0291
	CAG	0.11 (292)	0.13 (385)	p=0.0291
Asn	AAU	0.81 (6929)	0.83 (5753)	p=0.0009
	AAC*	0.19 (1653)	0.17 (1194)	p=0.0009
Lys	AAA	0.80 (8759)	0.82 (7441)	p=0.0088
	AAG	0.20 (2185)	0.18 (1689)	p=0.0088
Cys	UGU*	0.84 (1094)	0.75 (1547)	p<0.0001
	UGC	0.16 (204)	0.25 (524)	p<0.0001
Arg	CGU	0.11 (356)	0.20 (604)	p<0.0001
	CGC	0.01 (33)	0.03 (102)	p<0.0001
	CGA	0.06 (211)	0.14 (430)	p<0.0001
	CGG	0.00 (7)	0.01 (31)	p<0.0001
	AGA*	0.70 (2337)	0.46 (1381)	p<0.0001
	AGG	0.11 (376)	0.14 (428)	p=0.0003
Ser	UCU	0.22 (1310)	0.26 (1475)	p<0.0001
	UCC	0.10 (577)	0.09 (511)	p=0.1875
	UCA	0.28 (1721)	0.30 (1734)	p=0.0654
	UCG	0.04 (213)	0.05 (299)	p<0.0001
	AGU*	0.30 (1826)	0.24 (1393)	p<0.0001
	AGC	0.07 (394)	0.06 (362)	p=0.5749
Thr	ACU	0.30 (1263)	0.30 (1563)	p=0.8440
	ACC*	0.16 (650)	0.10 (516)	p<0.0001
	ACA	0.46 (1931)	0.50 (2584)	p=0.0009
	ACG	0.07 (311)	0.10 (511)	p<0.0001
Tyr	UAU	0.84 (3218)	0.83 (3708)	p=0.5355
	UAC	0.16 (613)	0.17 (733)	p=0.5355
Ala	GCU	0.45 (1598)	0.46 (2320)	p=0.2903
	GCC*	0.14 (477)	0.10 (505)	p<0.0001
	GCA	0.39 (1363)	0.36 (1819)	p=0.0375
	GCG	0.02 (87)	0.07 (346)	p<0.0001
Ile	AUU	0.43 (2816)	0.42 (3512)	p=0.0800
	AUC*	0.09 (613)	0.07 (575)	p<0.0001
	AUA	0.47 (3060)	0.51 (4282)	p<0.0001
Asp	GAU*	0.86 (4994)	0.81 (3763)	p<0.0001
	GAC	0.14 (819)	0.19 (878)	p<0.0001
Glu	GAA*	0.89 (6991)	0.86 (5363)	p<0.0001
	GAG	0.11 (892)	0.14 (905)	p<0.0001
Gly	GGU	0.42 (1763)	0.47 (1763)	p<0.0001
	GGC	0.03 (139)	0.04 (148)	p=0.1220
	GGA*	0.50 (2092)	0.38 (1434)	p<0.0001
	GGG	0.05 (203)	0.10 (391)	p<0.0001

Table 62 | Codon Usage from Expression Data for the Alveolate *P. falciparum*

Amino Acids	Codons	Top 5% Expressed Genes	Bottom 5% Expressed Genes	Chi-Sqaured
Phe	UUU	0.50 (2055)	0.55 (4483)	p<0.0001
	UUC*	0.50 (2095)	0.45 (3655)	p<0.0001
Leu	UUA	0.15 (1316)	0.16 (2507)	p=0.6584
	UUG	0.21 (1804)	0.21 (3442)	p=0.5506
	CUU	0.08 (671)	0.09 (1467)	p=0.0007
	CUC*	0.19 (1595)	0.17 (2777)	p<0.0001
	CUA	0.12 (1026)	0.13 (2023)	p=0.1785
	CUG	0.25 (2115)	0.24 (3805)	p=0.0662
Val	GUU	0.19 (1198)	0.22 (2206)	p=0.0002
	GUC*	0.22 (1326)	0.19 (1951)	p=0.0007
	GUA	0.19 (1194)	0.19 (1911)	p=0.4877
	GUG	0.40 (2444)	0.40 (4024)	p=0.7900
Pro	CCU	0.13 (539)	0.16 (1053)	p=0.0026
	CCC	0.39 (1553)	0.41 (2790)	p=0.0094
	CCA*	0.31 (1250)	0.25 (1704)	p<0.0001
	CCG	0.17 (685)	0.18 (1242)	p=0.0916
His	CAU	0.35 (873)	0.39 (2071)	p=0.0005
	CAC*	0.65 (1611)	0.61 (3202)	p=0.0005
Gln	CAA	0.46 (1715)	0.47 (2937)	p=0.7569
	CAG	0.54 (1975)	0.53 (3339)	p=0.7569
Asn	AAU	0.44 (3314)	0.48 (7031)	p<0.0001
	AAC*	0.56 (4155)	0.52 (7521)	p<0.0001
Lys	AAA	0.48 (5526)	0.53 (8550)	p<0.0001
	AAG*	0.52 (5997)	0.47 (7707)	p<0.0001
Cys	UGU	0.34 (606)	0.36 (1285)	p=0.1192
	UGC	0.66 (1192)	0.64 (2299)	p=0.1192
Arg	CGU	0.05 (223)	0.06 (482)	p=0.4268
	CGC	0.11 (444)	0.12 (1053)	p=0.0034
	CGA*	0.11 (472)	0.09 (790)	p=0.0005
	CGG	0.08 (321)	0.10 (898)	p<0.0001
	AGA*	0.34 (1431)	0.27 (2318)	p<0.0001
	AGG	0.31 (1323)	0.35 (3015)	p<0.0001
Ser	UCU	0.10 (820)	0.09 (1415)	p=0.2283
	UCC	0.21 (1774)	0.21 (3174)	p=0.5920
	UCA*	0.11 (878)	0.09 (1318)	p<0.0001
	UCG	0.12 (1028)	0.12 (1825)	p=0.5490
	AGU	0.16 (1359)	0.18 (2772)	p<0.0001
	AGC	0.30 (2458)	0.30 (4588)	p=0.1767
Thr	ACU	0.16 (815)	0.15 (1235)	p=0.0369
	ACC	0.33 (1681)	0.34 (2900)	p=0.0850
	ACA	0.21 (1097)	0.22 (1829)	p=0.7942
	ACG	0.30 (1514)	0.29 (2476)	p=0.7021
Tyr	UAU	0.34 (1340)	0.38 (2882)	p=0.0003
	UAC*	0.66 (2566)	0.62 (4763)	p=0.0003
Ala	GCU	0.16 (1222)	0.16 (1576)	p=0.8205
	GCC	0.28 (2124)	0.32 (3074)	p<0.0001
	GCA*	0.32 (2393)	0.26 (2565)	p<0.0001
	GCG	0.23 (1716)	0.26 (2476)	p=0.0001
Ile	AUU	0.38 (2166)	0.38 (4407)	p=0.2583
	AUC*	0.34 (1967)	0.30 (3445)	p<0.0001
	AUA	0.28 (1638)	0.32 (3619)	p<0.0001
Asp	GAU	0.46 (2987)	0.47 (4882)	p=0.2286
	GAC	0.54 (3442)	0.53 (5414)	p=0.2286
Glu	GAA	0.55 (5210)	0.54 (7556)	p=0.1304
	GAG	0.45 (4205)	0.46 (6351)	p=0.1304
Gly	GGU	0.17 (1280)	0.16 (1949)	p=0.0483
	GGC	0.24 (1843)	0.24 (2987)	p=0.9205
	GGA*	0.33 (2519)	0.26 (3240)	p<0.0001
	GGG	0.25 (1916)	0.33 (4105)	p<0.0001

Table 63 | Codon Usage from Expression Data for the Alveolate *P. vivax*

Amino Acids	Codons	Top 5% Expressed Genes	Bottom 5% Expressed Genes	Chi-Sqaured	Amino Acids	Codons	Top 5% Expressed Genes	Bottom 5% Expressed Genes	Chi-Sqaured
Phe	UUU	0.21 (1797)	0.41 (4480)	p<0.0001	**Ser**	UCU	0.10 (1723)	0.14 (3128)	p<0.0001
	UUC*	0.79 (6798)	0.59 (6476)	p<0.0001		UCC*	0.17 (2785)	0.15 (3211)	p<0.0001
Leu	UUA	0.02 (330)	0.08 (2126)	p<0.0001		UCA	0.07 (1155)	0.14 (2986)	p<0.0001
	UUG	0.13 (2437)	0.18 (4641)	p<0.0001		UCG*	0.34 (5623)	0.23 (5071)	p<0.0001
	CUU	0.13 (2347)	0.17 (4360)	p<0.0001		AGU	0.10 (1650)	0.16 (3584)	p<0.0001
	CUC*	0.24 (4474)	0.20 (5203)	p<0.0001		AGC*	0.21 (3520)	0.19 (4118)	p<0.0001
	CUA	0.06 (1153)	0.12 (3060)	p<0.0001	**Thr**	ACU	0.17 (2469)	0.23 (4053)	p<0.0001
	CUG*	0.42 (7895)	0.24 (6025)	p<0.0001		ACC*	0.28 (3992)	0.22 (3837)	p<0.0001
Val	GUU	0.13 (2207)	0.21 (4118)	p<0.0001		ACA	0.11 (1603)	0.23 (4005)	p<0.0001
	GUC	0.27 (4715)	0.28 (5448)	p=0.2611		ACG*	0.44 (6414)	0.33 (5785)	p<0.0001
	GUA	0.06 (1065)	0.15 (2959)	p<0.0001	**Tyr**	UAU	0.14 (969)	0.32 (2790)	p<0.0001
	GUG*	0.54 (9431)	0.37 (7221)	p<0.0001		UAC*	0.86 (5776)	0.68 (5887)	p<0.0001
Pro	CCU	0.23 (2476)	0.24 (2918)	p=0.0149	**Ala**	GCU	0.29 (6486)	0.30 (6927)	p=0.0083
	CCC*	0.30 (3249)	0.19 (2332)	p<0.0001		GCC*	0.40 (8854)	0.23 (5191)	p<0.0001
	CCA	0.12 (1310)	0.27 (3255)	p<0.0001		GCA	0.10 (2229)	0.23 (5199)	p<0.0001
	CCG*	0.35 (3848)	0.30 (3596)	p<0.0001		GCG	0.21 (4559)	0.24 (5430)	p<0.0001
His	CAU	0.16 (765)	0.37 (2570)	p<0.0001	**Ile**	AUU	0.29 (3105)	0.38 (4917)	p<0.0001
	CAC*	0.84 (4057)	0.63 (4308)	p<0.0001		AUC*	0.68 (7204)	0.50 (6414)	p<0.0001
Gln	CAA	0.20 (1708)	0.48 (5232)	p<0.0001		AUA	0.02 (254)	0.11 (1461)	p<0.0001
	CAG*	0.80 (6800)	0.52 (5637)	p<0.0001	**Asp**	GAU	0.23 (3094)	0.41 (6368)	p<0.0001
Asn	AAU	0.17 (1424)	0.38 (4213)	p<0.0001		GAC*	0.77 (10350)	0.59 (9007)	p<0.0001
	AAC*	0.83 (7066)	0.62 (6774)	p<0.0001	**Glu**	GAA	0.23 (3533)	0.45 (7360)	p<0.0001
Lys	AAA	0.13 (1921)	0.38 (5645)	p<0.0001		GAG*	0.77 (11734)	0.55 (9077)	p<0.0001
	AAG*	0.87 (13074)	0.62 (9094)	p<0.0001	**Gly**	GGU	0.29 (4851)	0.28 (4816)	p=0.2848
Cys	UGU	0.26 (935)	0.46 (2328)	p<0.0001		GGC*	0.52 (8814)	0.31 (5394)	p<0.0001
	UGC*	0.74 (2625)	0.54 (2693)	p<0.0001		GGA	0.14 (2425)	0.26 (4413)	p<0.0001
Arg	CGU*	0.36 (4205)	0.23 (3996)	p<0.0001		GGG	0.05 (895)	0.15 (2553)	p<0.0001
	CGC*	0.42 (4805)	0.23 (4003)	p<0.0001					
	CGA	0.08 (936)	0.23 (4103)	p<0.0001					
	CGG	0.07 (767)	0.12 (2087)	p<0.0001					
	AGA	0.04 (463)	0.12 (2091)	p<0.0001					
	AGG	0.03 (384)	0.07 (1310)	p<0.0001					

Table 64 | Codon Usage from Expression Data for the Stramenopile *P. infestans*

Amino Acids	Codons	Top 5% Expressed Genes	Bottom 5% Expressed Genes	Chi-Sqaured	Amino Acids	Codons	Top 5% Expressed Genes	Bottom 5% Expressed Genes	Chi-Sqaured
Phe	UUU	0.08 (956)	0.29 (4342)	p<0.0001	**Ser**	UCU	0.05 (1018)	0.12 (4575)	p<0.0001
	UUC*	0.92 (11218)	0.71 (10527)	p<0.0001		UCC*	0.20 (4331)	0.18 (6784)	p<0.0001
Leu	UUA	0.01 (146)	0.04 (1494)	p<0.0001		UCA	0.03 (561)	0.10 (3631)	p<0.0001
	UUG	0.06 (1464)	0.17 (6599)	p<0.0001		UCG*	0.40 (8656)	0.27 (10189)	p<0.0001
	CUU	0.05 (1405)	0.12 (4623)	p<0.0001		AGU	0.05 (976)	0.11 (4208)	p<0.0001
	CUC*	0.31 (8019)	0.27 (10462)	p<0.0001		AGC*	0.28 (6134)	0.22 (8491)	p<0.0001
	CUA	0.02 (433)	0.08 (3047)	p<0.0001	**Thr**	ACU	0.08 (1522)	0.19 (5067)	p<0.0001
	CUG*	0.55 (14195)	0.31 (11829)	p<0.0001		ACC*	0.32 (6058)	0.27 (7164)	p<0.0001
Val	GUU	0.06 (1456)	0.17 (5307)	p<0.0001		ACA	0.03 (652)	0.14 (3645)	p<0.0001
	GUC	0.30 (7034)	0.33 (10418)	p<0.0001		ACG*	0.56 (10500)	0.40 (10422)	p<0.0001
	GUA	0.02 (440)	0.09 (2740)	p<0.0001	**Tyr**	UAU	0.05 (445)	0.25 (2733)	p<0.0001
	GUG*	0.62 (14614)	0.41 (12858)	p<0.0001		UAC*	0.95 (8795)	0.75 (8399)	p<0.0001
Pro	CCU	0.11 (1567)	0.21 (4947)	p<0.0001	**Ala**	GCU	0.15 (4552)	0.25 (10698)	p<0.0001
	CCC*	0.41 (6130)	0.24 (5442)	p<0.0001		GCC*	0.53 (16542)	0.29 (12665)	p<0.0001
	CCA	0.05 (682)	0.20 (4543)	p<0.0001		GCA	0.04 (1268)	0.16 (6832)	p<0.0001
	CCG*	0.43 (6406)	0.35 (8178)	p<0.0001		GCG	0.28 (8648)	0.30 (12890)	p<0.0001
His	CAU	0.06 (393)	0.30 (2951)	p<0.0001	**Ile**	AUU	0.15 (2151)	0.29 (4761)	p<0.0001
	CAC*	0.94 (6193)	0.70 (6834)	p<0.0001		AUC*	0.84 (11906)	0.62 (10148)	p<0.0001
Gln	CAA	0.09 (1005)	0.37 (6032)	p<0.0001		AUA	0.01 (106)	0.09 (1410)	p<0.0001
	CAG*	0.91 (10596)	0.63 (10391)	p<0.0001	**Asp**	GAU	0.10 (1902)	0.32 (8320)	p<0.0001
Asn	AAU	0.06 (708)	0.26 (3686)	p<0.0001		GAC*	0.90 (16309)	0.68 (17624)	p<0.0001
	AAC*	0.94 (10979)	0.74 (10398)	p<0.0001	**Glu**	GAA	0.10 (1996)	0.33 (8600)	p<0.0001
Lys	AAA	0.04 (842)	0.24 (4652)	p<0.0001		GAG*	0.90 (18633)	0.67 (17546)	p<0.0001
	AAG*	0.96 (18422)	0.76 (14552)	p<0.0001	**Gly**	GGU	0.15 (3455)	0.21 (5781)	p<0.0001
Cys	UGU	0.12 (619)	0.34 (2396)	p<0.0001		GGC*	0.72 (16812)	0.40 (11050)	p<0.0001
	UGC*	0.88 (4758)	0.66 (4692)	p<0.0001		GGA	0.08 (1850)	0.20 (5570)	p<0.0001
Arg	CGU	0.17 (2623)	0.18 (5536)	p=0.4146		GGG	0.05 (1238)	0.19 (5239)	p<0.0001
	CGC*	0.65 (9866)	0.31 (9900)	p<0.0001					
	CGA	0.05 (721)	0.17 (5434)	p<0.0001					
	CGG	0.07 (1094)	0.17 (5245)	p<0.0001					
	AGA	0.02 (317)	0.08 (2375)	p<0.0001					
	AGG	0.04 (553)	0.09 (2978)	p<0.0001					

Table 65 | Codon Usage from Expression Data for the Stramenopile *P. sojae*

Amino Acids	Codons	Top 5% Expressed Genes	Bottom 5% Expressed Genes	Chi-Sqaured
Phe	UUU	0.64 (4483)	0.61 (2657)	p=0.0168
	UUC	0.36 (2568)	0.39 (1674)	p=0.0168
Leu	UUA	0.10 (1518)	0.13 (1355)	p<0.0001
	UUG	0.22 (3389)	0.22 (2235)	p=0.3955
	CUU*	0.25 (3833)	0.22 (2252)	p<0.0001
	CUC	0.14 (2146)	0.15 (1572)	p=0.0058
	CUA	0.11 (1633)	0.12 (1196)	p=0.0187
	CUG*	0.18 (2783)	0.16 (1690)	p=0.0002
Val	GUU	0.26 (3595)	0.28 (2018)	p=0.0174
	GUC	0.22 (3019)	0.23 (1669)	p=0.1241
	GUA	0.16 (2257)	0.19 (1352)	p<0.0001
	GUG*	0.35 (4878)	0.31 (2254)	p<0.0001
Pro	CCU*	0.31 (2817)	0.28 (1330)	p<0.0001
	CCC	0.17 (1482)	0.15 (717)	p=0.0130
	CCA	0.31 (2736)	0.35 (1703)	p<0.0001
	CCG	0.21 (1922)	0.22 (1056)	p=0.4847
His	CAU	0.43 (1739)	0.56 (1574)	p<0.0001
	CAC*	0.57 (2316)	0.44 (1212)	p<0.0001
Gln	CAA	0.47 (3254)	0.58 (2681)	p<0.0001
	CAG*	0.53 (3633)	0.42 (1910)	p<0.0001
Asn	AAU	0.52 (4151)	0.58 (2633)	p<0.0001
	AAC*	0.48 (3787)	0.42 (1930)	p<0.0001
Lys	AAA	0.39 (4697)	0.48 (3145)	p<0.0001
	AAG*	0.61 (7286)	0.52 (3415)	p<0.0001
Cys	UGU	0.50 (1711)	0.56 (1140)	p<0.0001
	UGC*	0.50 (1720)	0.44 (911)	p<0.0001
Arg	CGU*	0.39 (3697)	0.23 (1580)	p<0.0001
	CGC*	0.30 (2823)	0.20 (1364)	p<0.0001
	CGA	0.15 (1383)	0.23 (1544)	p<0.0001
	CGG	0.06 (569)	0.11 (772)	p<0.0001
	AGA	0.06 (596)	0.15 (1033)	p<0.0001
	AGG	0.04 (332)	0.08 (545)	p<0.0001
Ser	UCU	0.17 (2454)	0.18 (1549)	p=0.6665
	UCC	0.11 (1610)	0.11 (938)	p=0.0817
	UCA	0.18 (2483)	0.20 (1800)	p<0.0001
	UCG*	0.23 (3223)	0.18 (1600)	p<0.0001
	AGU	0.15 (2145)	0.17 (1513)	p<0.0001
	AGC	0.16 (2248)	0.16 (1427)	p=0.5542
Thr	ACU	0.27 (3358)	0.28 (1797)	p=0.8532
	ACC	0.17 (2099)	0.17 (1105)	p=0.7278
	ACA	0.25 (3025)	0.29 (1923)	p<0.0001
	ACG*	0.31 (3773)	0.26 (1703)	p<0.0001
Tyr	UAU	0.41 (2326)	0.52 (1624)	p<0.0001
	UAC*	0.59 (3380)	0.48 (1519)	p<0.0001
Ala	GCU*	0.39 (6314)	0.32 (2577)	p<0.0001
	GCC*	0.23 (3674)	0.17 (1384)	p<0.0001
	GCA	0.21 (3468)	0.29 (2387)	p<0.0001
	GCG	0.17 (2827)	0.22 (1828)	p<0.0001
Ile	AUU*	0.55 (5046)	0.47 (2585)	p<0.0001
	AUC	0.36 (3341)	0.34 (1879)	p=0.0041
	AUA	0.09 (859)	0.20 (1094)	p<0.0001
Asp	GAU	0.49 (5503)	0.56 (3474)	p<0.0001
	GAC*	0.51 (5621)	0.44 (2725)	p<0.0001
Glu	GAA	0.49 (5828)	0.53 (3415)	p<0.0001
	GAG*	0.51 (6056)	0.47 (2983)	p<0.0001
Gly	GGU*	0.37 (4997)	0.33 (1954)	p<0.0001
	GGC*	0.32 (4223)	0.25 (1517)	p<0.0001
	GGA	0.23 (3100)	0.28 (1689)	p<0.0001
	GGG	0.08 (1011)	0.14 (840)	p<0.0001

Table 66 | Codon Usage from Expression Data for the Stramenopile *P. halstedii*

Amino Acids	Codons	Top 5% Expressed Genes	Bottom 5% Expressed Genes	Chi-Sqaured
Phe	UUU	0.41 (3350)	0.43 (3455)	p=0.0083
	UUC	0.59 (4834)	0.57 (4584)	p=0.0083
Leu	UUA	0.02 (356)	0.04 (813)	p<0.0001
	UUG*	0.33 (6071)	0.25 (5132)	p<0.0001
	CUU	0.08 (1511)	0.12 (2451)	p<0.0001
	CUC*	0.28 (5111)	0.26 (5438)	p<0.0001
	CUA	0.03 (596)	0.07 (1455)	p<0.0001
	CUG	0.25 (4551)	0.26 (5417)	p=0.0095
Val	GUU	0.08 (1324)	0.15 (2423)	p<0.0001
	GUC	0.35 (5669)	0.34 (5644)	p=0.2906
	GUA	0.05 (854)	0.09 (1413)	p<0.0001
	GUG*	0.52 (8565)	0.43 (7125)	p<0.0001
Pro	CCU	0.17 (1697)	0.19 (2358)	p<0.0001
	CCC*	0.37 (3779)	0.31 (3751)	p<0.0001
	CCA	0.14 (1416)	0.23 (2854)	p<0.0001
	CCG*	0.32 (3194)	0.27 (3269)	p<0.0001
His	CAU	0.31 (1593)	0.35 (2323)	p<0.0001
	CAC*	0.69 (3619)	0.65 (4282)	p<0.0001
Gln	CAA	0.55 (4358)	0.56 (4862)	p=0.1284
	CAG	0.45 (3636)	0.44 (3869)	p=0.1284
Asn	AAU	0.14 (1155)	0.25 (2033)	p<0.0001
	AAC*	0.86 (7084)	0.75 (6123)	p<0.0001
Lys	AAA	0.19 (2309)	0.28 (3170)	p<0.0001
	AAG*	0.81 (9994)	0.72 (8128)	p<0.0001
Cys	UGU	0.23 (902)	0.36 (1411)	p<0.0001
	UGC*	0.77 (3050)	0.64 (2559)	p<0.0001
Arg	CGU	0.15 (1505)	0.19 (2630)	p<0.0001
	CGC*	0.53 (5330)	0.33 (4560)	p<0.0001
	CGA	0.14 (1370)	0.20 (2693)	p<0.0001
	CGG	0.14 (1426)	0.17 (2289)	p<0.0001
	AGA	0.02 (214)	0.06 (759)	p<0.0001
	AGG	0.02 (245)	0.06 (765)	p<0.0001
Ser	UCU	0.09 (1341)	0.11 (1733)	p<0.0001
	UCC	0.23 (3433)	0.22 (3522)	p=0.0377
	UCA	0.05 (803)	0.11 (1744)	p<0.0001
	UCG*	0.36 (5514)	0.27 (4467)	p<0.0001
	AGU	0.07 (1089)	0.10 (1652)	p<0.0001
	AGC	0.20 (3028)	0.20 (3186)	p=0.4103
Thr	ACU	0.10 (1345)	0.15 (2271)	p<0.0001
	ACC*	0.36 (4945)	0.33 (5074)	p<0.0001
	ACA	0.10 (1346)	0.17 (2639)	p<0.0001
	ACG*	0.44 (6082)	0.35 (5297)	p<0.0001
Tyr	UAU	0.16 (992)	0.24 (1478)	p<0.0001
	UAC*	0.84 (5140)	0.76 (4620)	p<0.0001
Ala	GCU	0.11 (2299)	0.17 (3489)	p<0.0001
	GCC*	0.43 (8688)	0.37 (7459)	p<0.0001
	GCA	0.10 (1944)	0.17 (3389)	p<0.0001
	GCG*	0.36 (7400)	0.30 (6040)	p<0.0001
Ile	AUU	0.29 (2776)	0.33 (2930)	p<0.0001
	AUC*	0.69 (6628)	0.60 (5385)	p<0.0001
	AUA	0.02 (212)	0.08 (699)	p<0.0001
Asp	GAU	0.19 (2512)	0.28 (3728)	p<0.0001
	GAC*	0.81 (10539)	0.72 (9541)	p<0.0001
Glu	GAA*	0.59 (6678)	0.50 (5806)	p<0.0001
	GAG	0.41 (4702)	0.50 (5853)	p<0.0001
Gly	GGU	0.17 (2349)	0.19 (2573)	p<0.0001
	GGC*	0.52 (7138)	0.42 (5619)	p<0.0001
	GGA	0.11 (1523)	0.17 (2317)	p<0.0001
	GGG	0.21 (2846)	0.22 (2906)	p=0.0231

Table 67 | Codon Usage from Expression Data for the Stramenopile *A. astaci*

Amino Acids	Codons	Top 5% Expressed Genes	Bottom 5% Expressed Genes	Chi-Sqaured
Phe	UUU	0.44 (3202)	0.42 (2478)	p=0.0618
	UUC	0.56 (4116)	0.58 (3403)	p=0.0618
Leu	UUA	0.02 (233)	0.03 (457)	p<0.0001
	UUG*	0.34 (5155)	0.22 (3314)	p<0.0001
	CUU	0.09 (1377)	0.14 (2059)	p<0.0001
	CUC*	0.32 (4751)	0.29 (4329)	p<0.0001
	CUA	0.02 (358)	0.06 (901)	p<0.0001
	CUG	0.21 (3149)	0.26 (3884)	p<0.0001
Val	GUU	0.12 (1576)	0.17 (2002)	p<0.0001
	GUC*	0.43 (5873)	0.39 (4707)	p<0.0001
	GUA	0.04 (612)	0.06 (781)	p<0.0001
	GUG*	0.41 (5591)	0.38 (4540)	p<0.0001
Pro	CCU	0.19 (1639)	0.20 (1628)	p=0.0890
	CCC*	0.33 (2865)	0.27 (2239)	p<0.0001
	CCA	0.14 (1246)	0.22 (1832)	p<0.0001
	CCG*	0.34 (2948)	0.30 (2491)	p<0.0001
His	CAU	0.27 (1123)	0.37 (1759)	p<0.0001
	CAC*	0.73 (2985)	0.63 (3047)	p<0.0001
Gln	CAA*	0.58 (3921)	0.54 (3477)	p<0.0001
	CAG	0.42 (2844)	0.46 (2989)	p<0.0001
Asn	AAU	0.15 (1061)	0.25 (1441)	p<0.0001
	AAC*	0.85 (6206)	0.75 (4217)	p<0.0001
Lys	AAA	0.20 (2197)	0.25 (2242)	p<0.0001
	AAG*	0.80 (8897)	0.75 (6595)	p<0.0001
Cys	UGU	0.18 (602)	0.33 (931)	p<0.0001
	UGC*	0.82 (2727)	0.67 (1905)	p<0.0001
Arg	CGU	0.18 (1572)	0.19 (2047)	p=0.0441
	CGC*	0.50 (4342)	0.35 (3699)	p<0.0001
	CGA	0.15 (1267)	0.18 (1925)	p<0.0001
	CGG	0.12 (1073)	0.16 (1718)	p<0.0001
	AGA	0.02 (172)	0.06 (610)	p<0.0001
	AGG	0.03 (243)	0.06 (623)	p<0.0001
Ser	UCU	0.09 (1072)	0.11 (1234)	p<0.0001
	UCC	0.19 (2316)	0.21 (2343)	p<0.0001
	UCA	0.05 (629)	0.10 (1082)	p<0.0001
	UCG*	0.41 (5058)	0.28 (3163)	p<0.0001
	AGU	0.07 (889)	0.10 (1137)	p<0.0001
	AGC	0.20 (2455)	0.20 (2246)	p=0.5950
Thr	ACU	0.11 (1312)	0.15 (1538)	p<0.0001
	ACC	0.31 (3591)	0.30 (3066)	p=0.0948
	ACA	0.11 (1315)	0.17 (1739)	p<0.0001
	ACG*	0.46 (5402)	0.38 (3924)	p<0.0001
Tyr	UAU	0.16 (829)	0.25 (1016)	p<0.0001
	UAC*	0.84 (4510)	0.75 (3013)	p<0.0001
Ala	GCU	0.15 (2665)	0.18 (2659)	p<0.0001
	GCC*	0.39 (6815)	0.35 (5286)	p<0.0001
	GCA	0.11 (2025)	0.17 (2579)	p<0.0001
	GCG*	0.35 (6148)	0.30 (4470)	p<0.0001
Ile	AUU	0.33 (2869)	0.32 (2069)	p=0.0312
	AUC*	0.65 (5558)	0.60 (3921)	p<0.0001
	AUA	0.02 (183)	0.08 (544)	p<0.0001
Asp	GAU	0.22 (2404)	0.28 (2892)	p<0.0001
	GAC*	0.78 (8523)	0.72 (7467)	p<0.0001
Glu	GAA*	0.62 (6208)	0.48 (4315)	p<0.0001
	GAG	0.38 (3797)	0.52 (4669)	p<0.0001
Gly	GGU	0.22 (2598)	0.21 (1928)	p=0.2510
	GGC*	0.48 (5768)	0.43 (3923)	p<0.0001
	GGA	0.11 (1316)	0.17 (1553)	p<0.0001
	GGG	0.19 (2318)	0.19 (1778)	p=0.9312

Table 68 | Codon Usage from Expression Data for the Stramenopile *A. invadans*

Amino Acids	Codons	Top 5% Expressed Genes	Bottom 5% Expressed Genes	Chi-Sqaured
Phe	UUU	0.39 (4072)	0.50 (8346)	p<0.0001
	UUC*	0.61 (6256)	0.50 (8256)	p<0.0001
Leu	UUA*	0.02 (385)	0.01 (390)	p<0.0001
	UUG	0.12 (2666)	0.13 (6098)	p=0.0019
	CUU	0.09 (2034)	0.10 (4771)	p=0.0002
	CUC*	0.64 (14302)	0.54 (25726)	p<0.0001
	CUA	0.01 (191)	0.02 (1169)	p<0.0001
	CUG	0.12 (2707)	0.20 (9490)	p<0.0001
Val	GUU	0.06 (1149)	0.08 (2625)	p<0.0001
	GUC*	0.64 (12214)	0.53 (18021)	p<0.0001
	GUA	0.02 (317)	0.03 (1037)	p<0.0001
	GUG	0.28 (5375)	0.36 (12186)	p<0.0001
Pro	CCU	0.07 (841)	0.07 (1689)	p=0.1551
	CCC*	0.33 (4206)	0.27 (6421)	p<0.0001
	CCA	0.06 (741)	0.13 (3243)	p<0.0001
	CCG	0.55 (6957)	0.53 (12803)	p=0.0037
His	CAU	0.14 (841)	0.21 (2696)	p<0.0001
	CAC*	0.86 (5290)	0.79 (10453)	p<0.0001
Gln	CAA	0.22 (1950)	0.34 (5739)	p<0.0001
	CAG*	0.78 (7038)	0.66 (10978)	p<0.0001
Asn	AAU	0.07 (685)	0.13 (1705)	p<0.0001
	AAC*	0.93 (9524)	0.87 (11620)	p<0.0001
Lys	AAA	0.06 (1087)	0.14 (2813)	p<0.0001
	AAG*	0.94 (15706)	0.86 (17512)	p<0.0001
Cys	UGU	0.08 (391)	0.14 (1197)	p<0.0001
	UGC*	0.92 (4443)	0.86 (7064)	p<0.0001
Arg	CGU*	0.13 (1697)	0.09 (2563)	p<0.0001
	CGC*	0.72 (9299)	0.59 (15906)	p<0.0001
	CGA	0.04 (474)	0.09 (2385)	p<0.0001
	CGG	0.08 (1079)	0.20 (5472)	p<0.0001
	AGA	0.02 (195)	0.01 (345)	p=0.0529
	AGG	0.01 (129)	0.01 (379)	p=0.0009
Ser	UCU	0.03 (516)	0.03 (1093)	p=0.0132
	UCC	0.12 (2144)	0.13 (4381)	p<0.0001
	UCA	0.02 (359)	0.04 (1199)	p<0.0001
	UCG*	0.58 (10064)	0.46 (15020)	p<0.0001
	AGU	0.03 (542)	0.06 (1884)	p<0.0001
	AGC	0.22 (3832)	0.27 (8888)	p<0.0001
Thr	ACU	0.04 (589)	0.04 (1168)	p=0.0081
	ACC	0.26 (4185)	0.26 (7268)	p=0.9216
	ACA	0.04 (663)	0.10 (2687)	p<0.0001
	ACG*	0.66 (10788)	0.61 (17101)	p<0.0001
Tyr	UAU	0.10 (762)	0.13 (1666)	p<0.0001
	UAC*	0.90 (7178)	0.87 (10691)	p<0.0001
Ala	GCU*	0.08 (2230)	0.07 (3563)	p<0.0001
	GCC*	0.49 (13136)	0.37 (17756)	p<0.0001
	GCA	0.04 (1187)	0.10 (4954)	p<0.0001
	GCG	0.39 (10437)	0.46 (22233)	p<0.0001
Ile	AUU	0.25 (3265)	0.30 (5500)	p<0.0001
	AUC*	0.74 (9646)	0.68 (12345)	p<0.0001
	AUA	0.02 (206)	0.02 (392)	p=0.0002
Asp	GAU	0.15 (2345)	0.16 (4215)	p=0.0263
	GAC	0.85 (13043)	0.84 (22029)	p=0.0263
Glu	GAA	0.29 (4291)	0.28 (6607)	p=0.4486
	GAG	0.71 (10737)	0.72 (16825)	p=0.4486
Gly	GGU*	0.20 (3649)	0.14 (3567)	p<0.0001
	GGC*	0.73 (13612)	0.67 (16556)	p<0.0001
	GGA	0.02 (340)	0.05 (1310)	p<0.0001
	GGG	0.05 (1003)	0.14 (3440)	p<0.0001

Table 69 | Codon Usage from Expression Data for the Stramenopile *S. parasitica*

Amino Acids	Codons	Top 5% Expressed Genes	Bottom 5% Expressed Genes	Chi-Sqaured	Amino Acids	Codons	Top 5% Expressed Genes	Bottom 5% Expressed Genes	Chi-Sqaured
Phe	UUU	0.36 (226)	0.50 (5568)	p<0.0001	**Ser**	UCU*	0.16 (208)	0.09 (2833)	p<0.0001
	UUC*	0.64 (402)	0.50 (5508)	p<0.0001		UCC	0.29 (364)	0.28 (8715)	p=0.1043
Leu	UUA*	0.03 (410)	0.02 (708)	p<0.0001		UCA*	0.08 (101)	0.06 (1881)	p<0.0001
	UUG*	0.26 (333)	0.12 (3554)	p<0.0001		UCG	0.20 (250)	0.24 (7552)	p<0.0001
	CUU*	0.26 (330)	0.10 (2933)	p<0.0001		AGU*	0.10 (132)	0.06 (1963)	p<0.0001
	CUC	0.26 (337)	0.35 (10241)	p<0.0001		AGC	0.16 (207)	0.26 (8105)	p<0.0001
	CUA	0.05 (647)	0.05 (1540)	p=0.3796	**Thr**	ACU*	0.21 (186)	0.10 (2043)	p<0.0001
	CUG	0.14 (181)	0.36 (10488)	p<0.0001		ACC	0.44 (393)	0.46 (9730)	p<0.0001
Val	GUU*	0.32 (359)	0.17 (3518)	p<0.0001		ACA	0.17 (151)	0.17 (3551)	p=0.7327
	GUC	0.39 (445)	0.43 (8897)	p<0.0001		ACG	0.18 (163)	0.28 (5876)	p<0.0001
	GUA*	0.12 (137)	0.08 (1730)	p<0.0001	**Tyr**	UAU	0.25 (102)	0.25 (1905)	p=0.7781
	GUG	0.17 (197)	0.32 (6614)	p<0.0001		UAC	0.75 (303)	0.75 (5703)	p=0.7781
Pro	CCU*	0.29 (205)	0.10 (2298)	p<0.0001	**Ala**	GCU*	0.23 (320)	0.11 (3661)	p<0.0001
	CCC	0.31 (218)	0.37 (8248)	p<0.0001		GCC	0.44 (608)	0.45 (14834)	p=0.0002
	CCA*	0.23 (163)	0.17 (3889)	p<0.0001		GCA*	0.19 (265)	0.17 (5514)	p<0.0001
	CCG	0.16 (113)	0.36 (7989)	p<0.0001		GCG	0.14 (201)	0.26 (8632)	p<0.0001
His	CAU*	0.37 (107)	0.21 (1845)	p<0.0001	**Ile**	AUU*	0.36 (268)	0.26 (3208)	p<0.0001
	CAC	0.63 (183)	0.79 (6787)	p<0.0001		AUC	0.58 (435)	0.62 (7631)	p<0.0001
Gln	CAA*	0.53 (270)	0.42 (4871)	p<0.0001		AUA	0.06 (485)	0.12 (1451)	p<0.0001
	CAG	0.47 (237)	0.58 (6766)	p<0.0001	**Asp**	GAU*	0.47 (468)	0.31 (7017)	p<0.0001
Asn	AAU*	0.31 (196)	0.26 (3581)	p<0.0001		GAC	0.53 (527)	0.69 (15308)	p<0.0001
	AAC	0.69 (429)	0.74 (10297)	p<0.0001	**Glu**	GAA*	0.59 (630)	0.46 (10961)	p<0.0001
Lys	AAA	0.30 (297)	0.38 (6466)	p<0.0001		GAG	0.41 (436)	0.54 (12950)	p<0.0001
	AAG*	0.70 (697)	0.62 (10732)	p<0.0001	**Gly**	GGU*	0.26 (339)	0.12 (3645)	p<0.0001
Cys	UGU*	0.32 (816)	0.24 (1309)	p<0.0001		GGC	0.23 (301)	0.35 (10447)	p<0.0001
	UGC	0.68 (170)	0.76 (4210)	p<0.0001		GGA*	0.40 (525)	0.23 (7024)	p<0.0001
Arg	CGU*	0.23 (175)	0.06 (1618)	p<0.0001		GGG	0.11 (142)	0.30 (9084)	p<0.0001
	CGC	0.19 (148)	0.21 (5636)	p=0.0002					
	CGA*	0.22 (170)	0.13 (3288)	p<0.0001					
	CGG	0.09 (661)	0.33 (8611)	p<0.0001					
	AGA*	0.16 (124)	0.09 (2477)	p<0.0001					
	AGG	0.10 (798)	0.18 (4630)	p<0.0001					

Table 70 | Codon Usage from Expression Data for the Stramenopile *P. multistriata*

Amino Acids	Codons	Top 5% Expressed Genes	Bottom 5% Expressed Genes	Chi-Sqaured	Amino Acids	Codons	Top 5% Expressed Genes	Bottom 5% Expressed Genes	Chi-Sqaured
Phe	UUU	0.51 (2571)	0.53 (7073)	p=0.0008	**Ser**	UCU	0.19 (2532)	0.21 (8118)	p<0.0001
	UUC*	0.49 (2520)	0.47 (6208)	p=0.0008		UCC*	0.18 (2447)	0.15 (5855)	p<0.0001
Leu	UUA*	0.09 (1076)	0.07 (2251)	p<0.0001		UCA	0.17 (2285)	0.18 (7131)	p=0.0064
	UUG	0.33 (3960)	0.33 (10883)	p=0.9953		UCG	0.14 (1924)	0.16 (6201)	p=0.0001
	CUU	0.19 (2261)	0.19 (6426)	p=0.1253		AGU	0.19 (2552)	0.18 (7133)	p=0.0148
	CUC*	0.21 (2575)	0.20 (6505)	p<0.0001		AGC	0.12 (1657)	0.13 (4959)	p=0.5087
	CUA	0.08 (995)	0.09 (3034)	p=0.0028	**Thr**	ACU	0.29 (2531)	0.30 (7393)	p=0.0576
	CUG	0.10 (1164)	0.12 (3962)	p<0.0001		ACC*	0.26 (2267)	0.21 (5148)	p<0.0001
Val	GUU	0.30 (2837)	0.29 (7315)	p=0.0731		ACA	0.25 (2223)	0.27 (6639)	p=0.0051
	GUC	0.24 (2286)	0.23 (5794)	p=0.0196		ACG	0.20 (1742)	0.22 (5495)	p<0.0001
	GUA	0.17 (1591)	0.17 (4406)	p=0.1533	**Tyr**	UAU	0.41 (1668)	0.43 (4351)	p=0.0091
	GUG	0.30 (2837)	0.31 (7948)	p=0.0064		UAC	0.59 (2445)	0.57 (5783)	p=0.0091
Pro	CCU	0.30 (2104)	0.31 (6182)	p=0.1313	**Ala**	GCU	0.32 (3612)	0.31 (9828)	p=0.3695
	CCC*	0.23 (1666)	0.19 (3806)	p<0.0001		GCC*	0.26 (2970)	0.22 (6844)	p<0.0001
	CCA	0.32 (2266)	0.34 (6880)	p=0.0011		GCA	0.30 (3392)	0.33 (10429)	p<0.0001
	CCG	0.15 (1062)	0.17 (3335)	p=0.0023		GCG	0.13 (1472)	0.14 (4497)	p=0.0003
His	CAU	0.49 (1457)	0.54 (4816)	p<0.0001	**Ile**	AUU	0.46 (3112)	0.46 (8406)	p=0.8325
	CAC*	0.51 (1491)	0.46 (4169)	p<0.0001		AUC*	0.42 (2810)	0.39 (7073)	p=0.0001
Gln	CAA	0.59 (2945)	0.56 (8899)	p=0.0035		AUA	0.12 (833)	0.15 (2708)	p<0.0001
	CAG	0.41 (2064)	0.44 (6865)	p=0.0035	**Asp**	GAU*	0.62 (5671)	0.60 (15377)	p<0.0001
Asn	AAU	0.47 (2881)	0.48 (8855)	p=0.2521		GAC	0.38 (3425)	0.40 (10259)	p<0.0001
	AAC	0.53 (3281)	0.52 (9750)	p=0.2521	**Glu**	GAA	0.45 (4407)	0.45 (11943)	p=0.4439
Lys	AAA	0.34 (2907)	0.35 (7413)	p=0.0164		GAG	0.55 (5394)	0.55 (14354)	p=0.4439
	AAG	0.66 (5731)	0.65 (13699)	p=0.0164	**Gly**	GGU	0.26 (2736)	0.26 (7218)	p=0.7747
Cys	UGU	0.54 (1449)	0.52 (3125)	p=0.0763		GGC	0.16 (1677)	0.18 (5024)	p<0.0001
	UGC	0.46 (1247)	0.48 (2920)	p=0.0763		GGA*	0.43 (4538)	0.40 (10990)	p<0.0001
Arg	CGU*	0.30 (2304)	0.27 (5664)	p<0.0001		GGG	0.15 (1544)	0.16 (4303)	p=0.0269
	CGC	0.13 (982)	0.12 (2553)	p=0.1640					
	CGA	0.16 (1267)	0.17 (3692)	p=0.0279					
	CGG	0.06 (447)	0.08 (1636)	p<0.0001					
	AGA	0.18 (1387)	0.19 (4096)	p=0.0050					
	AGG	0.18 (1357)	0.17 (3502)	p=0.0534					

Table 71 | Codon Usage from Expression Data for the Stramenopile *T. pseudonana*

Amino Acids	Codons	Top 5% Expressed Genes	Bottom 5% Expressed Genes	Chi-Sqaured
Phe	UUU*	0.45 (2687)	0.40 (2805)	p<0.0001
	UUC	0.55 (3285)	0.60 (4220)	p<0.0001
Leu	UUA*	0.08 (1109)	0.05 (1115)	p<0.0001
	UUG*	0.23 (3030)	0.21 (4698)	p=0.0003
	CUU*	0.17 (2266)	0.14 (3145)	p<0.0001
	CUC	0.20 (2643)	0.27 (6022)	p<0.0001
	CUA	0.06 (802)	0.06 (1243)	p=0.0891
	CUG	0.25 (3269)	0.26 (5687)	p=0.0311
Val	GUU*	0.17 (1690)	0.12 (1727)	p<0.0001
	GUC	0.29 (2870)	0.34 (4899)	p<0.0001
	GUA*	0.13 (1324)	0.10 (1474)	p<0.0001
	GUG	0.41 (4167)	0.44 (6251)	p=0.0011
Pro	CCU*	0.29 (2156)	0.27 (4127)	p<0.0001
	CCC	0.35 (2554)	0.36 (5655)	p=0.0160
	CCA	0.15 (1101)	0.15 (2303)	p=0.7465
	CCG	0.21 (1520)	0.22 (3418)	p=0.0244
His	CAU*	0.45 (1525)	0.39 (2284)	p<0.0001
	CAC	0.55 (1855)	0.61 (3567)	p<0.0001
Gln	CAA	0.44 (2361)	0.41 (3392)	p=0.0011
	CAG	0.56 (2964)	0.59 (4784)	p=0.0011
Asn	AAU	0.41 (2077)	0.42 (2289)	p=0.1113
	AAC	0.59 (3024)	0.58 (3129)	p=0.1113
Lys	AAA	0.36 (3104)	0.41 (3631)	p<0.0001
	AAG*	0.64 (5502)	0.59 (5188)	p<0.0001
Cys	UGU*	0.42 (1051)	0.36 (1306)	p<0.0001
	UGC	0.58 (1479)	0.64 (2358)	p<0.0001
Arg	CGU*	0.17 (1511)	0.12 (2170)	p<0.0001
	CGC*	0.26 (2256)	0.22 (3902)	p<0.0001
	CGA*	0.16 (1439)	0.14 (2521)	p<0.0001
	CGG	0.17 (1517)	0.25 (4325)	p<0.0001
	AGA*	0.11 (972)	0.09 (1566)	p<0.0001
	AGG	0.13 (1124)	0.17 (2919)	p<0.0001
Ser	UCU*	0.18 (1831)	0.14 (2628)	p<0.0001
	UCC	0.24 (2474)	0.30 (5641)	p<0.0001
	UCA*	0.12 (1254)	0.09 (1736)	p<0.0001
	UCG	0.18 (1840)	0.17 (3165)	p=0.0281
	AGU*	0.10 (1074)	0.09 (1728)	p=0.0007
	AGC	0.19 (1951)	0.22 (4115)	p<0.0001
Thr	ACU*	0.19 (1569)	0.15 (1699)	p<0.0001
	ACC	0.30 (2420)	0.33 (3899)	p<0.0001
	ACA	0.21 (1657)	0.19 (2264)	p=0.0450
	ACG	0.30 (2433)	0.33 (3836)	p<0.0001
Tyr	UAU	0.34 (1280)	0.33 (1501)	p=0.4360
	UAC	0.66 (2485)	0.67 (3022)	p=0.4360
Ala	GCU*	0.21 (2616)	0.16 (3500)	p<0.0001
	GCC	0.35 (4361)	0.39 (8365)	p<0.0001
	GCA	0.19 (2353)	0.17 (3724)	p=0.0011
	GCG	0.26 (3306)	0.28 (6037)	p=0.0004
Ile	AUU*	0.36 (2415)	0.30 (2047)	p<0.0001
	AUC	0.50 (3316)	0.55 (3785)	p<0.0001
	AUA	0.14 (914)	0.16 (1105)	p=0.0004
Asp	GAU*	0.38 (2657)	0.35 (3526)	p<0.0001
	GAC	0.62 (4270)	0.65 (6676)	p<0.0001
Glu	GAA	0.36 (3147)	0.34 (5403)	p=0.0042
	GAG	0.64 (5545)	0.66 (10312)	p=0.0042
Gly	GGU*	0.20 (2247)	0.14 (2725)	p<0.0001
	GGC*	0.33 (3665)	0.30 (5963)	p<0.0001
	GGA	0.23 (2529)	0.23 (4527)	p=0.8781
	GGG	0.24 (2733)	0.34 (6854)	p<0.0001

Table 72 | Codon Usage from Expression Data for the Stramenopile *N. gaditana*

Amino Acids	Codons	Top 5% Expressed Genes	Bottom 5% Expressed Genes	Chi-Sqaured	Amino Acids	Codons	Top 5% Expressed Genes	Bottom 5% Expressed Genes	Chi-Sqaured
Phe	UUU*	0.17 (588)	0.11 (1318)	p<0.0001	**Ser**	UCU*	0.07 (386)	0.03 (724)	p<0.0001
	UUC	0.83 (2785)	0.89 (11118)	p<0.0001		UCC	0.30 (1792)	0.29 (6441)	p=0.0245
Leu	UUA	0.03 (257)	0.02 (618)	p<0.0001		UCA*	0.06 (377)	0.03 (636)	p<0.0001
	UUG	0.07 (607)	0.07 (2325)	p=0.0421		UCG	0.32 (1906)	0.38 (8482)	p<0.0001
	CUU*	0.07 (574)	0.04 (1239)	p<0.0001		AGU*	0.05 (302)	0.03 (581)	p<0.0001
	CUC	0.60 (4914)	0.59 (20110)	p=0.0228		AGC	0.19 (1138)	0.24 (5446)	p<0.0001
	CUA*	0.03 (283)	0.02 (731)	p<0.0001	**Thr**	ACU*	0.07 (343)	0.04 (659)	p<0.0001
	CUG	0.19 (1522)	0.27 (9141)	p<0.0001		ACC	0.24 (1178)	0.22 (3743)	p=0.0019
Val	GUU*	0.07 (442)	0.03 (904)	p<0.0001		ACA*	0.08 (392)	0.04 (740)	p<0.0001
	GUC	0.62 (4096)	0.66 (17587)	p<0.0001		ACG	0.62 (3084)	0.70 (12258)	p<0.0001
	GUA*	0.04 (290)	0.02 (542)	p<0.0001	**Tyr**	UAU*	0.14 (345)	0.08 (689)	p<0.0001
	GUG	0.27 (1811)	0.29 (7734)	p=0.0091		UAC	0.86 (2105)	0.92 (7819)	p<0.0001
Pro	CCU*	0.08 (410)	0.04 (892)	p<0.0001	**Ala**	GCU*	0.06 (692)	0.03 (1741)	p<0.0001
	CCC*	0.40 (2087)	0.37 (8524)	p<0.0001		GCC*	0.38 (4562)	0.34 (20257)	p<0.0001
	CCA*	0.08 (411)	0.04 (875)	p<0.0001		GCA*	0.07 (800)	0.03 (1482)	p<0.0001
	CCG	0.44 (2299)	0.55 (12724)	p<0.0001		GCG	0.50 (6069)	0.60 (35565)	p<0.0001
His	CAU*	0.20 (442)	0.12 (931)	p<0.0001	**Ile**	AUU*	0.13 (473)	0.10 (851)	p<0.0001
	CAC	0.80 (1758)	0.88 (6903)	p<0.0001		AUC	0.79 (2967)	0.84 (7019)	p<0.0001
Gln	CAA*	0.18 (589)	0.15 (1164)	p<0.0001		AUA*	0.08 (301)	0.06 (475)	p<0.0001
	CAG	0.82 (2667)	0.85 (6725)	p<0.0001	**Asp**	GAU*	0.14 (965)	0.07 (1961)	p<0.0001
Asn	AAU*	0.18 (503)	0.13 (950)	p<0.0001		GAC	0.86 (5862)	0.93 (24392)	p<0.0001
	AAC	0.82 (2242)	0.87 (6595)	p<0.0001	**Glu**	GAA*	0.15 (982)	0.11 (2253)	p<0.0001
Lys	AAA	0.13 (786)	0.12 (1418)	p=0.0049		GAG	0.85 (5697)	0.89 (18367)	p<0.0001
	AAG	0.87 (5192)	0.88 (10709)	p=0.0049	**Gly**	GGU*	0.07 (530)	0.04 (1110)	p<0.0001
Cys	UGU*	0.13 (246)	0.07 (522)	p<0.0001		GGC	0.77 (6001)	0.77 (23365)	p=0.9643
	UGC	0.87 (1714)	0.93 (6951)	p<0.0001		GGA*	0.05 (420)	0.04 (1217)	p<0.0001
Arg	CGU*	0.05 (363)	0.04 (1282)	p<0.0001		GGG	0.11 (849)	0.15 (4668)	p<0.0001
	CGC*	0.57 (4051)	0.53 (17073)	p<0.0001					
	CGA	0.06 (445)	0.05 (1741)	p=0.0043					
	CGG	0.22 (1554)	0.32 (10285)	p<0.0001					
	AGA*	0.05 (325)	0.02 (486)	p<0.0001					
	AGG*	0.05 (331)	0.04 (1172)	p<0.0001					

Table 73 | Codon Usage from Expression Data for the Stramenopile *A. anophagefferens*

Amino Acids	Codons	Top 5% Expressed Genes	Bottom 5% Expressed Genes	Chi-Sqaured
Phe	UUU	0.36 (4891)	0.48 (8292)	p<0.0001
	UUC*	0.64 (8531)	0.52 (9147)	p<0.0001
Leu	UUA	0.09 (2539)	0.11 (3922)	p<0.0001
	UUG	0.14 (4097)	0.20 (7609)	p<0.0001
	CUU*	0.25 (7452)	0.20 (7577)	p<0.0001
	CUC*	0.24 (7092)	0.20 (7434)	p<0.0001
	CUA*	0.13 (3789)	0.11 (4086)	p<0.0001
	CUG	0.15 (4513)	0.18 (6682)	p<0.0001
Val	GUU*	0.32 (7413)	0.28 (7395)	p<0.0001
	GUC	0.25 (5808)	0.24 (6249)	p=0.0030
	GUA*	0.17 (4027)	0.16 (4062)	p<0.0001
	GUG	0.26 (5914)	0.32 (8416)	p<0.0001
Pro	CCU*	0.39 (6798)	0.31 (5994)	p<0.0001
	CCC	0.18 (3238)	0.19 (3802)	p=0.0149
	CCA	0.29 (5021)	0.32 (6227)	p<0.0001
	CCG	0.14 (2507)	0.18 (3547)	p<0.0001
His	CAU	0.42 (3156)	0.49 (5487)	p<0.0001
	CAC*	0.58 (4379)	0.51 (5771)	p<0.0001
Gln	CAA	0.40 (5548)	0.54 (8523)	p<0.0001
	CAG*	0.60 (8235)	0.46 (7249)	p<0.0001
Asn	AAU	0.39 (5912)	0.50 (10403)	p<0.0001
	AAC*	0.61 (9436)	0.50 (10295)	p<0.0001
Lys	AAA	0.37 (10859)	0.52 (15055)	p<0.0001
	AAG*	0.63 (18439)	0.48 (14078)	p<0.0001
Cys	UGU	0.34 (1895)	0.40 (3577)	p<0.0001
	UGC*	0.66 (3714)	0.60 (5326)	p<0.0001
Arg	CGU*	0.19 (4024)	0.13 (3439)	p<0.0001
	CGC*	0.19 (4028)	0.15 (3968)	p<0.0001
	CGA	0.16 (3246)	0.17 (4363)	p=0.0005
	CGG	0.08 (1582)	0.11 (2757)	p<0.0001
	AGA	0.19 (3851)	0.25 (6390)	p<0.0001
	AGG	0.19 (3953)	0.19 (4897)	p=0.7002

Amino Acids	Codons	Top 5% Expressed Genes	Bottom 5% Expressed Genes	Chi-Sqaured
Ser	UCU*	0.21 (5346)	0.18 (6766)	p<0.0001
	UCC*	0.21 (5295)	0.15 (5678)	p<0.0001
	UCA	0.15 (3842)	0.19 (7034)	p<0.0001
	UCG	0.14 (3600)	0.14 (5313)	p=0.2832
	AGU	0.11 (2707)	0.14 (5297)	p<0.0001
	AGC	0.19 (4886)	0.19 (7001)	p=0.6300
Thr	ACU*	0.26 (5250)	0.24 (5842)	p<0.0001
	ACC*	0.33 (6632)	0.23 (5709)	p<0.0001
	ACA	0.23 (4698)	0.31 (7590)	p<0.0001
	ACG	0.19 (3775)	0.22 (5413)	p<0.0001
Tyr	UAU	0.34 (3679)	0.46 (4729)	p<0.0001
	UAC*	0.66 (7277)	0.54 (5603)	p<0.0001
Ala	GCU*	0.37 (11031)	0.29 (8684)	p<0.0001
	GCC	0.18 (5472)	0.20 (5984)	p<0.0001
	GCA	0.30 (9031)	0.33 (9884)	p<0.0001
	GCG	0.14 (4264)	0.19 (5723)	p<0.0001
Ile	AUU	0.36 (6873)	0.39 (8581)	p<0.0001
	AUC*	0.45 (8497)	0.38 (8329)	p<0.0001
	AUA	0.19 (3571)	0.23 (5111)	p<0.0001
Asp	GAU	0.52 (10869)	0.56 (12643)	p<0.0001
	GAC*	0.48 (10164)	0.44 (10123)	p<0.0001
Glu	GAA	0.46 (10912)	0.52 (13972)	p<0.0001
	GAG*	0.54 (12830)	0.48 (12689)	p<0.0001
Gly	GGU*	0.21 (5615)	0.20 (5273)	p=0.0008
	GGC	0.24 (6277)	0.24 (6362)	p=0.2213
	GGA*	0.38 (9991)	0.34 (8809)	p<0.0001
	GGG	0.17 (4483)	0.22 (5777)	p<0.0001

Table 74 | Codon Usage from Expression Data for the Rhizarian *B. natans*

Amino Acids	Codons	Top 5% Expressed Genes	Bottom 5% Expressed Genes	Chi-Sqaured
Phe	UUU	0.11 (499)	0.15 (1069)	p<0.0001
	UUC*	0.89 (3916)	0.85 (6015)	p<0.0001
Leu	UUA	0.01 (97)	0.03 (506)	p<0.0001
	UUG	0.10 (1027)	0.13 (2143)	p<0.0001
	CUU	0.09 (967)	0.09 (1453)	p=0.1032
	CUC	0.22 (2335)	0.20 (3475)	p=0.0018
	CUA	0.02 (256)	0.06 (957)	p<0.0001
	CUG*	0.56 (5900)	0.50 (8427)	p<0.0001
Val	GUU	0.16 (1626)	0.15 (2219)	p=0.0571
	GUC*	0.57 (5732)	0.51 (7446)	p<0.0001
	GUA	0.03 (317)	0.06 (882)	p<0.0001
	GUG	0.24 (2399)	0.28 (4008)	p<0.0001
Pro	CCU	0.14 (834)	0.14 (1293)	p=0.8886
	CCC	0.23 (1367)	0.25 (2272)	p=0.0325
	CCA	0.10 (621)	0.18 (1655)	p<0.0001
	CCG*	0.52 (3094)	0.43 (4005)	p<0.0001
His	CAU	0.28 (790)	0.37 (1585)	p<0.0001
	CAC*	0.72 (2015)	0.63 (2744)	p<0.0001
Gln	CAA	0.15 (757)	0.24 (1686)	p<0.0001
	CAG*	0.85 (4168)	0.76 (5360)	p<0.0001
Asn	AAU	0.15 (658)	0.22 (1316)	p<0.0001
	AAC*	0.85 (3713)	0.78 (4744)	p<0.0001
Lys	AAA	0.14 (1153)	0.27 (1609)	p<0.0001
	AAG*	0.86 (7071)	0.73 (4427)	p<0.0001
Cys	UGU	0.18 (315)	0.24 (794)	p<0.0001
	UGC*	0.82 (1400)	0.76 (2544)	p<0.0001
Arg	CGU*	0.22 (1754)	0.14 (1772)	p<0.0001
	CGC*	0.43 (3384)	0.31 (3811)	p<0.0001
	CGA	0.10 (783)	0.16 (1963)	p<0.0001
	CGG	0.17 (1326)	0.22 (2723)	p<0.0001
	AGA	0.02 (141)	0.05 (622)	p<0.0001
	AGG	0.06 (505)	0.12 (1463)	p<0.0001
Ser	UCU	0.07 (560)	0.07 (936)	p=0.5118
	UCC	0.22 (1883)	0.20 (2808)	p=0.0024
	UCA	0.05 (425)	0.09 (1171)	p<0.0001
	UCG*	0.41 (3506)	0.36 (4979)	p<0.0001
	AGU	0.04 (360)	0.06 (827)	p<0.0001
	AGC	0.21 (1746)	0.22 (2981)	p=0.0393
Thr	ACU	0.08 (546)	0.08 (925)	p=0.0674
	ACC	0.30 (2106)	0.31 (3344)	p=0.2004
	ACA	0.07 (488)	0.13 (1454)	p<0.0001
	ACG*	0.56 (3971)	0.48 (5237)	p<0.0001
Tyr	UAU	0.14 (464)	0.20 (907)	p<0.0001
	UAC*	0.86 (2814)	0.80 (3641)	p<0.0001
Ala	GCU	0.12 (1532)	0.12 (2261)	p=0.5612
	GCC*	0.43 (5536)	0.40 (7394)	p<0.0001
	GCA	0.08 (983)	0.13 (2429)	p<0.0001
	GCG*	0.37 (4784)	0.35 (6520)	p<0.0001
Ile	AUU	0.22 (1391)	0.21 (1816)	p=0.0186
	AUC*	0.75 (4673)	0.71 (6256)	p<0.0001
	AUA	0.03 (201)	0.08 (739)	p<0.0001
Asp	GAU	0.26 (2098)	0.29 (3014)	p<0.0001
	GAC*	0.74 (5995)	0.71 (7447)	p<0.0001
Glu	GAA	0.28 (2180)	0.36 (2779)	p<0.0001
	GAG*	0.72 (5633)	0.64 (4869)	p<0.0001
Gly	GGU*	0.19 (1747)	0.14 (1638)	p<0.0001
	GGC*	0.54 (5031)	0.49 (5605)	p<0.0001
	GGA	0.09 (864)	0.12 (1357)	p<0.0001
	GGG	0.19 (1759)	0.25 (2874)	p<0.0001

Table 75 | Codon Usage from Expression Data for the Rhizarian *P. brassicae*

6.4 Codon Usage – Protein Coding Domains and Non-Domains

General description of tables 76 - 107: The codon usage data is given in two forms, inside and outside of brackets. The numbers inside of the brackets represent the amount the codon was found within the genome files. The numbers outside the brackets represent the proportion of that the codon was found within the set of genes as opposed to synonymous codons. The p-values shown were calculated by a 2x2 contingency table, any p-value above 0.001 was deemed insignificant, these are shown using red text.

Amino Acids	Codons	Domain	Non-Domain	Chi-Sqaured
Phe	UUU	0.04 (48)	0.13 (35)	p<0.0001
	UUC*	0.96 (1061)	0.87 (242)	p<0.0001
Leu	UUA	0.00 (1)	0.01 (4)	p=0.0024
	UUG	0.04 (93)	0.06 (38)	p=0.0881
	CUU	0.08 (170)	0.09 (54)	p=0.6528
	CUC	0.60 (1284)	0.54 (340)	p=0.0035
	CUA	0.00 (4)	0.01 (8)	p=0.0003
	CUG	0.27 (583)	0.30 (190)	p=0.1896
Val	GUU	0.07 (118)	0.14 (69)	p<0.0001
	GUC	0.53 (937)	0.48 (243)	p=0.0569
	GUA	0.01 (17)	0.02 (10)	p=0.0619
	GUG	0.40 (702)	0.36 (184)	p=0.1916
Pro	CCU	0.10 (109)	0.14 (61)	p=0.0382
	CCC	0.69 (749)	0.66 (294)	p=0.2687
	CCA	0.00 (3)	0.03 (13)	p<0.0001
	CCG	0.21 (232)	0.18 (80)	p=0.1351
His	CAU	0.07 (39)	0.11 (21)	p=0.0301
	CAC	0.93 (554)	0.89 (162)	p=0.0301
Gln	CAA	0.03 (29)	0.13 (42)	p<0.0001
	CAG*	0.97 (812)	0.87 (288)	p<0.0001
Asn	AAU	0.05 (48)	0.11 (31)	p<0.0001
	AAC*	0.95 (976)	0.89 (252)	p<0.0001
Lys	AAA	0.01 (24)	0.03 (28)	p=0.0004
	AAG*	0.99 (1783)	0.97 (796)	p=0.0004
Cys	UGU	0.08 (23)	0.21 (14)	p=0.0014
	UGC	0.92 (268)	0.79 (52)	p=0.0014
Arg	CGU*	0.17 (209)	0.10 (45)	p=0.0007
	CGC	0.47 (579)	0.38 (169)	p=0.0015
	CGA	0.03 (42)	0.08 (36)	p<0.0001
	CGG	0.03 (42)	0.07 (29)	p=0.0047
	AGA	0.02 (26)	0.05 (22)	p=0.0019
	AGG	0.27 (327)	0.31 (138)	p=0.0575
Ser	UCU	0.06 (78)	0.10 (60)	p=0.0043
	UCC	0.33 (413)	0.33 (199)	p=0.9971
	UCA	0.01 (11)	0.03 (16)	p=0.0028
	UCG*	0.37 (463)	0.27 (161)	p<0.0001
	AGU	0.01 (11)	0.03 (16)	p=0.0028
	AGC	0.22 (282)	0.25 (154)	p=0.1523
Thr	ACU	0.11 (154)	0.15 (68)	p=0.0387
	ACC	0.71 (967)	0.65 (295)	p=0.0161
	ACA	0.01 (12)	0.05 (24)	p<0.0001
	ACG	0.17 (232)	0.15 (68)	p=0.3071
Tyr	UAU	0.07 (59)	0.13 (28)	p=0.0053
	UAC	0.93 (754)	0.87 (183)	p=0.0053
Ala	GCU	0.16 (313)	0.22 (232)	p<0.0001
	GCC*	0.71 (1369)	0.64 (673)	p=0.0003
	GCA	0.01 (22)	0.03 (30)	p=0.0006
	GCG	0.12 (229)	0.11 (111)	p=0.3116
Ile	AUU	0.13 (163)	0.13 (35)	p=0.9709
	AUC	0.87 (1102)	0.84 (229)	p=0.1456
	AUA	0.00 (3)	0.04 (10)	p<0.0001
Asp	GAU	0.24 (318)	0.27 (128)	p=0.1834
	GAC	0.76 (1025)	0.73 (351)	p=0.1834
Glu	GAA	0.06 (82)	0.11 (69)	p<0.0001
	GAG*	0.94 (1327)	0.89 (547)	p<0.0001
Gly	GGU	0.31 (529)	0.28 (161)	p=0.1722
	GGC	0.63 (1074)	0.57 (330)	p=0.0172
	GGA	0.05 (80)	0.10 (58)	p<0.0001
	GGG	0.01 (25)	0.05 (27)	p<0.0001

Table 76 | Codon Usage from Domains and Non-Domains for the Amoebozoan *A. castellani*

Amino Acids	Codons	Domain	Non-Domain	Chi-Sqaured	Amino Acids	Codons	Domain	Non-Domain	Chi-Sqaured
Phe	UUU	0.41 (104)	0.38 (51)	p=0.6632	**Ser**	UCU	0.20 (84)	0.27 (62)	p=0.0389
	UUC	0.59 (152)	0.62 (82)	p=0.6632		UCC	0.05 (21)	0.05 (12)	p=0.8924
Leu	UUA	0.29 (151)	0.26 (59)	p=0.4433		UCA	0.50 (214)	0.40 (93)	p=0.0123
	UUG	0.08 (41)	0.09 (20)	p=0.6409		UCG	0.01 (3)	0.00 (0)	p=0.2001
	CUU	0.55 (292)	0.51 (116)	p=0.2625		AGU	0.17 (71)	0.24 (55)	p=0.0284
	CUC	0.08 (43)	0.12 (27)	p=0.1074		AGC	0.07 (31)	0.04 (9)	p=0.0811
	CUA	0.00 (1)	0.01 (2)	p=0.1674	**Thr**	ACU	0.62 (256)	0.62 (152)	p=0.9887
	CUG	0.00 (0)	0.02 (4)	p=0.0023		ACC	0.13 (55)	0.13 (33)	p=0.9558
Val	GUU	0.74 (409)	0.73 (191)	p=0.8420		ACA	0.24 (99)	0.24 (58)	p=0.9310
	GUC	0.15 (85)	0.11 (28)	p=0.0752		ACG	0.01 (3)	0.01 (2)	p=0.8978
	GUA	0.10 (58)	0.15 (38)	p=0.0913	**Tyr**	UAU	0.76 (172)	0.80 (94)	p=0.4146
	GUG	0.01 (4)	0.02 (5)	p=0.1282		UAC	0.24 (55)	0.20 (24)	p=0.4146
Pro	CCU	0.04 (9)	0.04 (8)	p=0.8015	**Ala**	GCU	0.68 (391)	0.67 (294)	p=0.6452
	CCC	0.00 (1)	0.03 (6)	p=0.0254		GCC	0.08 (47)	0.14 (62)	p=0.0027
	CCA	0.96 (236)	0.92 (179)	p=0.0990		GCA	0.24 (137)	0.19 (86)	p=0.0980
	CCG	0.00 (0)	0.01 (1)	p=0.2596		GCG	0.00 (1)	0.00 (0)	p=0.3808
His	CAU	0.75 (117)	0.72 (41)	p=0.6504	**Ile**	AUU*	0.81 (426)	0.69 (122)	p=0.0003
	CAC	0.25 (39)	0.28 (16)	p=0.6504		AUC	0.16 (83)	0.25 (44)	p=0.0081
Gln	CAA	0.99 (230)	1.00 (160)	p=0.1496		AUA	0.03 (14)	0.07 (12)	p=0.0132
	CAG	0.01 (3)	0.00 (0)	p=0.1496	**Asp**	GAU	0.81 (258)	0.77 (138)	p=0.3501
Asn	AAU	0.75 (300)	0.71 (149)	p=0.3289		GAC	0.19 (62)	0.23 (41)	p=0.3501
	AAC	0.25 (102)	0.29 (61)	p=0.3289	**Glu**	GAA	0.93 (350)	0.93 (301)	p=0.8363
Lys	AAA	0.65 (548)	0.57 (364)	p=0.0043		GAG	0.07 (26)	0.07 (21)	p=0.8363
	AAG	0.35 (299)	0.43 (270)	p=0.0043	**Gly**	GGU	0.30 (149)	0.36 (80)	p=0.1232
Cys	UGU	0.74 (136)	0.71 (109)	p=0.5276		GGC	0.01 (5)	0.05 (11)	p=0.0010
	UGC	0.26 (47)	0.29 (44)	p=0.5276		GGA	0.67 (331)	0.58 (128)	p=0.0144
Arg	CGU	0.11 (46)	0.07 (8)	p=0.3313		GGG	0.02 (8)	0.01 (3)	p=0.7850
	CGC	0.00 (0)	0.00 (0)	-					
	CGA	0.02 (7)	0.03 (3)	p=0.4149					
	CGG	0.00 (0)	0.00 (0)	-					
	AGA	0.87 (380)	0.89 (96)	p=0.5887					
	AGG	0.01 (4)	0.01 (1)	p=0.9917					

Table 77 | Codon Usage from Domains and Non-Domains for the Amoebozoan *E. histolytica*

Amino Acids	Codons	Domain	Non-Domain	Chi-Sqaured	Amino Acids	Codons	Domain	Non-Domain	Chi-Sqaured
Phe	UUU	0.34 (110)	0.51 (83)	p=0.0005	Ser	UCU	0.24 (125)	0.17 (91)	p=0.0045
	UUC*	0.66 (212)	0.49 (81)	p=0.0005		UCC	0.12 (60)	0.08 (42)	p=0.0423
Leu	UUA	0.58 (334)	0.65 (184)	p=0.0491		UCA	0.47 (239)	0.52 (275)	p=0.0761
	UUG	0.07 (41)	0.07 (19)	p=0.8267		UCG	0.01 (4)	0.01 (4)	p=0.9673
	CUU	0.13 (73)	0.11 (32)	p=0.5654		AGU	0.12 (61)	0.14 (73)	p=0.3515
	CUC	0.22 (129)	0.15 (43)	p=0.0132		AGC	0.04 (23)	0.08 (42)	p=0.0207
	CUA	0.01 (3)	0.02 (7)	p=0.0122	Thr	ACU	0.44 (213)	0.40 (242)	p=0.1924
	CUG	0.00 (0)	0.00 (0)	-		ACC	0.43 (209)	0.35 (211)	p=0.0056
Val	GUU	0.56 (310)	0.60 (193)	p=0.2784		ACA	0.13 (62)	0.25 (151)	p<0.0001
	GUC	0.32 (176)	0.22 (72)	p=0.0025		ACG	0.00 (1)	0.00 (1)	p=0.8754
	GUA	0.11 (62)	0.15 (49)	p=0.0883	Tyr	UAU	0.44 (99)	0.61 (82)	p=0.0018
	GUG	0.00 (2)	0.02 (7)	p=0.0105		UAC	0.56 (127)	0.39 (53)	p=0.0018
Pro	CCU	0.02 (7)	0.05 (36)	p=0.0347	Ala	GCU	0.47 (253)	0.49 (252)	p=0.4798
	CCC	0.01 (3)	0.01 (4)	p=0.4896		GCC*	0.36 (195)	0.26 (134)	p=0.0004
	CCA	0.97 (307)	0.94 (669)	p=0.0629		GCA	0.17 (92)	0.25 (126)	p=0.0027
	CCG	0.00 (0)	0.00 (2)	p=0.3445		GCG	0.00 (0)	0.00 (2)	p=0.1468
His	CAU	0.42 (53)	0.60 (69)	p=0.0064	Ile	AUU	0.48 (224)	0.57 (145)	p=0.0192
	CAC	0.58 (72)	0.40 (46)	p=0.0064		AUC*	0.49 (228)	0.29 (74)	p<0.0001
Gln	CAA	1.00 (296)	0.98 (615)	p=0.0573		AUA	0.03 (13)	0.13 (34)	p<0.0001
	CAG	0.00 (1)	0.02 (12)	p=0.0573	Asp	GAU	0.81 (290)	0.79 (202)	p=0.5356
Asn	AAU	0.46 (190)	0.69 (352)	p<0.0001		GAC	0.19 (67)	0.21 (53)	p=0.5356
	AAC*	0.54 (226)	0.31 (159)	p<0.0001	Glu	GAA	0.92 (398)	0.94 (419)	p=0.1918
Lys	AAA	0.75 (508)	0.77 (443)	p=0.4458		GAG	0.08 (35)	0.06 (26)	p=0.1918
	AAG	0.25 (170)	0.23 (134)	p=0.4458	Gly	GGU*	0.92 (388)	0.76 (265)	p<0.0001
Cys	UGU	0.74 (64)	0.82 (32)	p=0.3489		GGC	0.03 (12)	0.07 (24)	p=0.0080
	UGC	0.26 (22)	0.18 (7)	p=0.3489		GGA	0.05 (20)	0.17 (58)	p<0.0001
Arg	CGU*	0.42 (149)	0.25 (51)	p<0.0001		GGG	0.00 (2)	0.00 (1)	p=0.6792
	CGC	0.00 (0)	0.00 (0)	-					
	CGA	0.00 (0)	0.00 (0)	-					
	CGG	0.00 (0)	0.00 (0)	-					
	AGA	0.57 (200)	0.75 (156)	p<0.0001					
	AGG	0.01 (2)	0.00 (0)	p=0.2766					

Table 78 | Codon Usage from Domains and Non-Domains for the Amoebozoan *D. discoideum*

Amino Acids	Codons	Domain	Non-Domain	Chi-Sqaured	Amino Acids	Codons	Domain	Non-Domain	Chi-Sqaured
Phe	UUU	0.30 (523)	0.31 (370)	p=0.5733	**Ser**	UCU	0.22 (708)	0.21 (688)	p=0.3169
	UUC	0.70 (1218)	0.69 (823)	p=0.5733		UCC	0.21 (656)	0.21 (680)	p=0.7238
Leu	UUA	0.10 (391)	0.10 (284)	p=0.4012		UCA	0.15 (465)	0.17 (545)	p=0.0156
	UUG	0.22 (904)	0.20 (566)	p=0.0691		UCG	0.14 (455)	0.14 (450)	p=0.6289
	CUU	0.07 (290)	0.09 (246)	p=0.0089		AGU	0.13 (418)	0.12 (395)	p=0.2505
	CUC	0.30 (1229)	0.28 (772)	p=0.0316		AGC	0.15 (472)	0.15 (476)	p=0.8638
	CUA	0.06 (261)	0.09 (245)	p=0.0002	**Thr**	ACU	0.20 (478)	0.19 (399)	p=0.5415
	CUG	0.25 (1011)	0.24 (676)	p=0.6328		ACC*	0.32 (771)	0.27 (554)	p<0.0001
Val	GUU	0.15 (429)	0.15 (268)	p=0.4858		ACA	0.23 (562)	0.28 (588)	p=0.0002
	GUC	0.41 (1157)	0.37 (679)	p=0.0027		ACG	0.25 (590)	0.26 (539)	p=0.3027
	GUA	0.13 (365)	0.15 (275)	p=0.0637	**Tyr**	UAU	0.34 (475)	0.41 (358)	p=0.0005
	GUG	0.30 (837)	0.33 (609)	p=0.0202		UAC*	0.66 (931)	0.59 (514)	p=0.0005
Pro	CCU	0.25 (422)	0.22 (363)	p=0.0273	**Ala**	GCU	0.22 (605)	0.21 (386)	p=0.2296
	CCC*	0.40 (667)	0.32 (534)	p<0.0001		GCC*	0.40 (1083)	0.32 (582)	p<0.0001
	CCA	0.23 (388)	0.31 (507)	p<0.0001		GCA	0.23 (609)	0.29 (530)	p<0.0001
	CCG	0.12 (199)	0.15 (251)	p=0.0054		GCG	0.15 (400)	0.19 (346)	p=0.0004
His	CAU	0.27 (282)	0.35 (265)	p=0.0003	**Ile**	AUU	0.16 (426)	0.20 (345)	p=0.0018
	CAC*	0.73 (747)	0.65 (484)	p=0.0003		AUC*	0.72 (1916)	0.66 (1154)	p<0.0001
Gln	CAA	0.43 (674)	0.48 (631)	p=0.0025		AUA	0.12 (319)	0.15 (257)	p=0.0106
	CAG	0.57 (910)	0.52 (679)	p=0.0025	**Asp**	GAU	0.51 (1066)	0.54 (950)	p=0.0241
Asn	AAU	0.37 (706)	0.41 (616)	p=0.0298		GAC	0.49 (1044)	0.46 (804)	p=0.0241
	AAC	0.63 (1205)	0.59 (902)	p=0.0298	**Glu**	GAA	0.40 (1063)	0.45 (997)	p=0.0010
Lys	AAA	0.31 (751)	0.34 (603)	p=0.0252		GAG*	0.60 (1583)	0.55 (1226)	p=0.0010
	AAG	0.69 (1666)	0.66 (1152)	p=0.0252	**Gly**	GGU	0.30 (831)	0.26 (465)	p=0.0033
Cys	UGU	0.46 (274)	0.55 (241)	p=0.0076		GGC	0.16 (444)	0.20 (353)	p=0.0013
	UGC	0.54 (317)	0.45 (199)	p=0.0076		GGA	0.44 (1216)	0.42 (737)	p=0.0696
Arg	CGU	0.11 (229)	0.10 (199)	p=0.2421		GGG	0.09 (252)	0.12 (217)	p=0.0010
	CGC	0.10 (211)	0.11 (223)	p=0.3284					
	CGA	0.17 (355)	0.17 (346)	p=0.9115					
	CGG	0.04 (74)	0.05 (98)	p=0.0363					
	AGA	0.40 (820)	0.36 (708)	p=0.0049					
	AGG	0.17 (356)	0.20 (404)	p=0.0145					

Table 79 | Codon Usage from Domains and Non-Domains for the Amoebozoan *P. fungivorum*

Amino Acids	Codons	Domain	Non-Domain	Chi-Sqaured
Phe	UUU	0.42 (341)	0.43 (199)	p=0.8905
	UUC	0.58 (465)	0.57 (267)	p=0.8905
Leu	UUA	0.00 (0)	0.00 (3)	p=0.0354
	UUG	0.02 (36)	0.06 (73)	p<0.0001
	CUU	0.13 (223)	0.18 (202)	p=0.0012
	CUC*	0.59 (1000)	0.50 (581)	p<0.0001
	CUA	0.01 (13)	0.02 (26)	p=0.0008
	CUG	0.25 (429)	0.23 (269)	p=0.2438
Val	GUU	0.12 (173)	0.16 (158)	p=0.0019
	GUC*	0.56 (825)	0.45 (445)	p<0.0001
	GUA	0.01 (9)	0.02 (16)	p=0.0136
	GUG	0.32 (477)	0.37 (368)	p=0.0083
Pro	CCU	0.10 (73)	0.11 (134)	p=0.2751
	CCC	0.22 (167)	0.19 (235)	p=0.1997
	CCA	0.04 (33)	0.10 (118)	p<0.0001
	CCG	0.64 (491)	0.60 (720)	p=0.0403
His	CAU	0.14 (53)	0.30 (83)	p<0.0001
	CAC*	0.86 (336)	0.70 (196)	p<0.0001
Gln	CAA	0.04 (33)	0.12 (87)	p<0.0001
	CAG*	0.96 (705)	0.88 (633)	p<0.0001
Asn	AAU	0.04 (30)	0.12 (54)	p<0.0001
	AAC*	0.96 (775)	0.88 (393)	p<0.0001
Lys	AAA	0.01 (11)	0.02 (18)	p=0.0285
	AAG	0.99 (1017)	0.98 (731)	p=0.0285
Cys	UGU	0.08 (30)	0.16 (48)	p=0.0013
	UGC	0.92 (340)	0.84 (249)	p=0.0013
Arg	CGU	0.13 (110)	0.13 (97)	p=0.8797
	CGC*	0.58 (498)	0.45 (335)	p<0.0001
	CGA	0.03 (22)	0.07 (54)	p<0.0001
	CGG	0.23 (198)	0.27 (197)	p=0.1076
	AGA	0.01 (9)	0.03 (21)	p=0.0088
	AGG	0.02 (20)	0.05 (37)	p=0.0043
Ser	UCU	0.07 (91)	0.09 (131)	p=0.0518
	UCC	0.14 (189)	0.16 (241)	p=0.1375
	UCA	0.05 (67)	0.07 (103)	p=0.0348
	UCG*	0.68 (922)	0.54 (817)	p<0.0001
	AGU	0.01 (15)	0.03 (41)	p=0.0019
	AGC	0.05 (64)	0.11 (172)	p<0.0001
Thr	ACU	0.09 (114)	0.13 (109)	p=0.0120
	ACC*	0.56 (680)	0.45 (379)	p<0.0001
	ACA	0.04 (48)	0.10 (87)	p<0.0001
	ACG	0.31 (375)	0.32 (273)	p=0.5062
Tyr	UAU	0.08 (46)	0.11 (26)	p=0.1184
	UAC	0.92 (555)	0.89 (210)	p=0.1184
Ala	GCU	0.16 (306)	0.22 (485)	p<0.0001
	GCC*	0.50 (975)	0.44 (955)	p<0.0001
	GCA	0.06 (123)	0.09 (197)	p=0.0010
	GCG	0.28 (550)	0.25 (541)	p=0.0160
Ile	AUU	0.21 (231)	0.30 (110)	p=0.0004
	AUC*	0.79 (863)	0.68 (247)	p<0.0001
	AUA	0.00 (5)	0.02 (8)	p=0.0022
Asp	GAU	0.18 (216)	0.24 (202)	p=0.0002
	GAC*	0.82 (1013)	0.76 (632)	p=0.0002
Glu	GAA	0.05 (57)	0.09 (67)	p=0.0001
	GAG*	0.95 (1208)	0.91 (699)	p=0.0001
Gly	GGU	0.17 (241)	0.17 (189)	p=0.9664
	GGC	0.68 (965)	0.67 (735)	p=0.3609
	GGA	0.05 (70)	0.08 (91)	p=0.0008
	GGG	0.10 (137)	0.08 (89)	p=0.1547

Table 80 | Codon Usage from Domains and Non-Domains for the Apusomonadan *T. trahens*

Amino Acids	Codons	Domain	Non-Domain	Chi-Sqaured
Phe	UUU	0.07 (60)	0.25 (82)	$p<0.0001$
	UUC*	0.93 (825)	0.75 (249)	$p<0.0001$
Leu	UUA	0.04 (76)	0.05 (39)	$p=0.3728$
	UUG	0.01 (23)	0.07 (56)	$p<0.0001$
	CUU	0.06 (112)	0.11 (82)	$p=0.0001$
	CUC	0.47 (840)	0.42 (322)	$p=0.0126$
	CUA	0.01 (10)	0.03 (20)	$p<0.0001$
	CUG*	0.40 (713)	0.32 (248)	$p=0.0002$
Val	GUU	0.13 (212)	0.19 (102)	$p=0.0007$
	GUC*	0.72 (1150)	0.55 (293)	$p<0.0001$
	GUA	0.01 (18)	0.03 (15)	$p=0.0058$
	GUG	0.13 (215)	0.22 (118)	$p<0.0001$
Pro	CCU	0.06 (51)	0.12 (77)	$p<0.0001$
	CCC*	0.66 (574)	0.53 (340)	$p<0.0001$
	CCA	0.01 (5)	0.05 (32)	$p<0.0001$
	CCG	0.27 (234)	0.29 (187)	$p=0.3232$
His	CAU	0.12 (55)	0.32 (62)	$p<0.0001$
	CAC*	0.88 (386)	0.68 (134)	$p<0.0001$
Gln	CAA	0.02 (13)	0.13 (37)	$p<0.0001$
	CAG*	0.98 (673)	0.87 (240)	$p<0.0001$
Asn	AAU	0.12 (104)	0.18 (59)	$p=0.0023$
	AAC	0.88 (800)	0.82 (265)	$p=0.0023$
Lys	AAA	0.04 (48)	0.13 (65)	$p<0.0001$
	AAG*	0.96 (1279)	0.87 (451)	$p<0.0001$
Cys	UGU	0.06 (17)	0.29 (58)	$p<0.0001$
	UGC*	0.94 (268)	0.71 (141)	$p<0.0001$
Arg	CGU	0.20 (243)	0.18 (97)	$p=0.2937$
	CGC*	0.73 (899)	0.58 (319)	$p<0.0001$
	CGA	0.00 (5)	0.04 (23)	$p<0.0001$
	CGG	0.05 (64)	0.12 (67)	$p<0.0001$
	AGA	0.02 (21)	0.05 (28)	$p<0.0001$
	AGG	0.01 (7)	0.04 (20)	$p<0.0001$

Amino Acids	Codons	Domain	Non-Domain	Chi-Sqaured
Ser	UCU	0.15 (205)	0.22 (157)	$p=0.0004$
	UCC*	0.58 (777)	0.44 (324)	$p<0.0001$
	UCA	0.01 (17)	0.02 (18)	$p=0.0433$
	UCG	0.09 (125)	0.13 (94)	$p=0.0120$
	AGU	0.02 (24)	0.02 (15)	$p=0.6718$
	AGC	0.14 (191)	0.17 (121)	$p=0.1566$
Thr	ACU	0.11 (138)	0.13 (60)	$p=0.2888$
	ACC*	0.84 (1033)	0.73 (333)	$p<0.0001$
	ACA	0.02 (22)	0.07 (31)	$p<0.0001$
	ACG	0.03 (33)	0.07 (33)	$p<0.0001$
Tyr	UAU	0.12 (75)	0.30 (49)	$p<0.0001$
	UAC*	0.88 (566)	0.70 (113)	$p<0.0001$
Ala	GCU	0.16 (303)	0.19 (206)	$p=0.0211$
	GCC*	0.79 (1524)	0.68 (742)	$p<0.0001$
	GCA	0.02 (31)	0.04 (41)	$p=0.0002$
	GCG	0.04 (80)	0.09 (101)	$p<0.0001$
Ile	AUU	0.21 (246)	0.19 (60)	$p=0.2841$
	AUC	0.76 (884)	0.72 (232)	$p=0.0807$
	AUA	0.02 (28)	0.10 (32)	$p<0.0001$
Asp	GAU	0.28 (308)	0.34 (137)	$p=0.0178$
	GAC	0.72 (809)	0.66 (268)	$p=0.0178$
Glu	GAA	0.02 (21)	0.10 (40)	$p<0.0001$
	GAG*	0.98 (1081)	0.90 (380)	$p<0.0001$
Gly	GGU*	0.43 (649)	0.30 (163)	$p<0.0001$
	GGC	0.52 (781)	0.53 (290)	$p=0.6940$
	GGA	0.01 (16)	0.06 (32)	$p<0.0001$
	GGG	0.04 (66)	0.12 (66)	$p<0.0001$

Table 81 | Codon Usage from Domains and Non-Domains for the Opisthokont *F. alba*

Amino Acids	Codons	Domain	Non-Domain	Chi-Sqaured
Phe	UUU	0.29 (266)	0.50 (192)	p<0.0001
	UUC*	0.71 (636)	0.50 (190)	p<0.0001
Leu	UUA	0.00 (6)	0.01 (4)	p=0.5411
	UUG	0.05 (96)	0.15 (115)	p<0.0001
	CUU	0.15 (266)	0.17 (134)	p=0.2530
	CUC*	0.59 (1023)	0.41 (320)	p<0.0001
	CUA	0.01 (13)	0.02 (14)	p=0.0190
	CUG	0.20 (344)	0.26 (201)	p=0.0009
Val	GUU	0.21 (331)	0.26 (181)	p=0.0114
	GUC*	0.66 (1036)	0.48 (339)	p<0.0001
	GUA	0.01 (10)	0.02 (16)	p=0.0006
	GUG	0.13 (202)	0.24 (167)	p<0.0001
Pro	CCU	0.10 (84)	0.18 (117)	p<0.0001
	CCC*	0.65 (529)	0.48 (302)	p<0.0001
	CCA	0.07 (53)	0.13 (81)	p<0.0001
	CCG	0.18 (142)	0.21 (135)	p=0.0776
His	CAU	0.10 (43)	0.29 (68)	p<0.0001
	CAC*	0.90 (395)	0.71 (170)	p<0.0001
Gln	CAA	0.22 (174)	0.31 (123)	p=0.0007
	CAG*	0.78 (614)	0.69 (272)	p=0.0007
Asn	AAU	0.09 (80)	0.17 (65)	p<0.0001
	AAC*	0.91 (802)	0.83 (321)	p<0.0001
Lys	AAA	0.03 (40)	0.10 (54)	p<0.0001
	AAG*	0.97 (1348)	0.90 (499)	p<0.0001
Cys	UGU	0.05 (20)	0.21 (33)	p<0.0001
	UGC*	0.95 (377)	0.79 (124)	p<0.0001
Arg	CGU	0.19 (217)	0.19 (102)	p=0.8880
	CGC*	0.71 (817)	0.56 (298)	p<0.0001
	CGA	0.06 (73)	0.10 (55)	p=0.0042
	CGG	0.02 (23)	0.05 (25)	p=0.0020
	AGA	0.00 (3)	0.03 (17)	p<0.0001
	AGG	0.01 (14)	0.06 (34)	p<0.0001
Ser	UCU	0.17 (219)	0.17 (157)	p=0.7741
	UCC	0.29 (382)	0.27 (248)	p=0.2979
	UCA	0.04 (50)	0.06 (54)	p=0.0215
	UCG	0.31 (402)	0.29 (262)	p=0.3066
	AGU	0.02 (20)	0.03 (31)	p=0.0038
	AGC	0.18 (233)	0.17 (159)	p=0.8140
Thr	ACU	0.22 (288)	0.24 (161)	p=0.5192
	ACC*	0.62 (800)	0.48 (327)	p<0.0001
	ACA	0.05 (58)	0.08 (52)	p=0.0040
	ACG	0.11 (142)	0.21 (141)	p<0.0001
Tyr	UAU	0.12 (81)	0.21 (50)	p=0.0010
	UAC*	0.88 (571)	0.79 (185)	p=0.0010
Ala	GCU	0.27 (501)	0.30 (475)	p=0.0558
	GCC*	0.58 (1069)	0.44 (686)	p<0.0001
	GCA	0.05 (98)	0.10 (159)	p<0.0001
	GCG	0.09 (160)	0.16 (243)	p<0.0001
Ile	AUU	0.37 (440)	0.45 (159)	p=0.0093
	AUC	0.63 (747)	0.54 (192)	p=0.0026
	AUA	0.00 (2)	0.01 (5)	p=0.0023
Asp	GAU	0.21 (241)	0.32 (138)	p<0.0001
	GAC*	0.79 (904)	0.68 (295)	p<0.0001
Glu	GAA	0.17 (191)	0.27 (112)	p<0.0001
	GAG*	0.83 (931)	0.73 (299)	p<0.0001
Gly	GGU	0.19 (296)	0.24 (149)	p=0.0104
	GGC*	0.76 (1156)	0.63 (384)	p<0.0001
	GGA	0.03 (53)	0.09 (56)	p<0.0001
	GGG	0.02 (23)	0.04 (23)	p=0.0012

Table 82 | Codon Usage from Domains and Non-Domains for the Opisthokont *C. owczarzaki*

Amino Acids	Codons	Domain	Non-Domain	Chi-Sqaured
Phe	UUU	0.19 (85)	0.36 (52)	p<0.0001
	UUC*	0.81 (353)	0.64 (94)	p<0.0001
Leu	UUA	0.06 (66)	0.09 (33)	p=0.0327
	UUG	0.19 (212)	0.19 (69)	p=0.9574
	CUU*	0.25 (273)	0.16 (57)	p=0.0006
	CUC	0.13 (147)	0.15 (54)	p=0.3871
	CUA	0.13 (141)	0.17 (62)	p=0.0289
	CUG	0.24 (259)	0.23 (80)	p=0.6834
Val	GUU	0.26 (272)	0.19 (63)	p=0.0085
	GUC	0.23 (241)	0.26 (88)	p=0.2131
	GUA	0.19 (197)	0.25 (83)	p=0.0169
	GUG	0.33 (350)	0.31 (104)	p=0.4419
Pro	CCU	0.25 (126)	0.35 (100)	p=0.0038
	CCC*	0.51 (257)	0.28 (82)	p<0.0001
	CCA	0.14 (70)	0.23 (65)	p=0.0019
	CCG	0.10 (50)	0.14 (41)	p=0.0685
His	CAU	0.19 (50)	0.39 (42)	p<0.0001
	CAC*	0.81 (208)	0.61 (66)	p<0.0001
Gln	CAA	0.24 (112)	0.37 (75)	p=0.0002
	CAG*	0.76 (364)	0.63 (126)	p=0.0002
Asn	AAU	0.26 (136)	0.45 (99)	p<0.0001
	AAC*	0.74 (387)	0.55 (122)	p<0.0001
Lys	AAA	0.16 (164)	0.24 (80)	p=0.0005
	AAG*	0.84 (894)	0.76 (257)	p=0.0005
Cys	UGU	0.45 (79)	0.67 (49)	p=0.0014
	UGC	0.55 (97)	0.33 (24)	p=0.0014
Arg	CGU*	0.42 (324)	0.31 (106)	p=0.0007
	CGC	0.25 (193)	0.21 (73)	p=0.2068
	CGA	0.13 (104)	0.12 (40)	p=0.4381
	CGG	0.04 (32)	0.08 (26)	p=0.0154
	AGA	0.11 (83)	0.20 (69)	p<0.0001
	AGG	0.05 (39)	0.08 (27)	p=0.0598
Ser	UCU	0.19 (158)	0.21 (84)	p=0.5627
	UCC	0.26 (211)	0.18 (73)	p=0.0024
	UCA	0.16 (127)	0.13 (52)	p=0.2081
	UCG	0.12 (100)	0.11 (44)	p=0.4826
	AGU	0.14 (111)	0.20 (80)	p=0.0052
	AGC	0.13 (110)	0.18 (72)	p=0.0462
Thr	ACU	0.27 (191)	0.25 (86)	p=0.5096
	ACC	0.40 (285)	0.30 (104)	p=0.0018
	ACA	0.19 (140)	0.26 (91)	p=0.0135
	ACG	0.14 (102)	0.19 (67)	p=0.0344
Tyr	UAU	0.21 (83)	0.32 (35)	p=0.0193
	UAC	0.79 (310)	0.68 (75)	p=0.0193
Ala	GCU	0.31 (369)	0.32 (160)	p=0.6976
	GCC	0.41 (494)	0.34 (170)	p=0.0038
	GCA	0.12 (146)	0.19 (96)	p=0.0002
	GCG	0.16 (187)	0.15 (77)	p=0.8650
Ile	AUU	0.33 (270)	0.32 (57)	p=0.7365
	AUC	0.57 (458)	0.49 (87)	p=0.0626
	AUA	0.10 (82)	0.19 (34)	p=0.0008
Asp	GAU	0.55 (484)	0.51 (161)	p=0.1757
	GAC	0.45 (390)	0.49 (155)	p=0.1757
Glu	GAA	0.25 (228)	0.35 (145)	p<0.0001
	GAG*	0.75 (695)	0.65 (272)	p<0.0001
Gly	GGU	0.44 (435)	0.42 (149)	p=0.5326
	GGC	0.20 (200)	0.18 (65)	p=0.4480
	GGA	0.30 (297)	0.29 (103)	p=0.7379
	GGG	0.05 (51)	0.10 (35)	p=0.0018

Table 83 | Codon Usage from Domains and Non-Domains for the Opisthokont *S. arctica*

Amino Acids	Codons	Domain	Non-Domain	Chi-Sqaured	Amino Acids	Codons	Domain	Non-Domain	Chi-Sqaured
Phe	UUU	0.33 (365)	0.44 (302)	$p<0.0001$	**Ser**	UCU	0.21 (500)	0.23 (488)	$p=0.0523$
	UUC*	0.67 (726)	0.56 (392)	$p<0.0001$		UCC	0.22 (538)	0.19 (400)	$p=0.0069$
Leu	UUA	0.03 (105)	0.05 (99)	$p<0.0001$		UCA	0.07 (178)	0.10 (217)	$p=0.0006$
	UUG	0.09 (341)	0.11 (222)	$p=0.0347$		UCG	0.15 (367)	0.13 (274)	$p=0.0378$
	CUU	0.25 (920)	0.27 (532)	$p=0.2299$		AGU	0.11 (272)	0.12 (247)	$p=0.6700$
	CUC*	0.37 (1333)	0.27 (540)	$p<0.0001$		AGC	0.24 (575)	0.23 (496)	$p=0.8463$
	CUA	0.06 (217)	0.08 (157)	$p=0.0062$	**Thr**	ACU	0.20 (461)	0.23 (377)	$p=0.0245$
	CUG	0.20 (721)	0.22 (435)	$p=0.0689$		ACC	0.28 (646)	0.26 (427)	$p=0.1880$
Val	GUU	0.25 (618)	0.31 (421)	$p<0.0001$		ACA	0.23 (526)	0.26 (418)	$p=0.0540$
	GUC*	0.46 (1124)	0.32 (425)	$p<0.0001$		ACG	0.29 (673)	0.25 (416)	$p=0.0097$
	GUA	0.07 (176)	0.11 (155)	$p<0.0001$	**Tyr**	UAU	0.37 (450)	0.43 (311)	$p=0.0108$
	GUG	0.22 (540)	0.26 (347)	$p=0.0095$		UAC	0.63 (765)	0.57 (413)	$p=0.0108$
Pro	CCU	0.24 (331)	0.29 (349)	$p=0.0054$	**Ala**	GCU	0.24 (875)	0.25 (586)	$p=0.4606$
	CCC*	0.30 (409)	0.23 (272)	$p<0.0001$		GCC	0.33 (1200)	0.31 (728)	$p=0.1075$
	CCA	0.18 (250)	0.24 (289)	$p<0.0001$		GCA	0.22 (823)	0.26 (624)	$p=0.0006$
	CCG	0.27 (370)	0.24 (281)	$p=0.0411$		GCG	0.21 (786)	0.19 (443)	$p=0.0098$
His	CAU	0.29 (231)	0.34 (158)	$p=0.0367$	**Ile**	AUU	0.28 (608)	0.32 (377)	$p=0.0209$
	CAC	0.71 (573)	0.66 (301)	$p=0.0367$		AUC*	0.54 (1189)	0.46 (545)	$p<0.0001$
Gln	CAA	0.22 (318)	0.30 (316)	$p<0.0001$		AUA	0.18 (395)	0.23 (272)	$p=0.0010$
	CAG*	0.78 (1099)	0.70 (753)	$p<0.0001$	**Asp**	GAU	0.37 (908)	0.45 (685)	$p<0.0001$
Asn	AAU	0.36 (627)	0.40 (417)	$p=0.0482$		GAC*	0.63 (1536)	0.55 (832)	$p<0.0001$
	AAC	0.64 (1121)	0.60 (636)	$p=0.0482$	**Glu**	GAA	0.22 (654)	0.29 (458)	$p<0.0001$
Lys	AAA	0.16 (400)	0.21 (282)	$p<0.0001$		GAG*	0.78 (2350)	0.71 (1129)	$p<0.0001$
	AAG*	0.84 (2133)	0.79 (1044)	$p<0.0001$	**Gly**	GGU	0.19 (521)	0.18 (261)	$p=0.6815$
Cys	UGU	0.32 (386)	0.36 (228)	$p=0.1180$		GGC	0.37 (1042)	0.35 (508)	$p=0.1792$
	UGC	0.68 (804)	0.64 (402)	$p=0.1180$		GGA	0.23 (645)	0.23 (328)	$p=0.8173$
Arg	CGU	0.18 (352)	0.17 (194)	$p=0.5251$		GGG	0.21 (602)	0.24 (355)	$p=0.0272$
	CGC*	0.39 (767)	0.28 (318)	$p<0.0001$					
	CGA	0.05 (105)	0.09 (100)	$p=0.0003$					
	CGG	0.07 (132)	0.10 (118)	$p=0.0005$					
	AGA	0.10 (203)	0.14 (160)	$p=0.0026$					
	AGG	0.20 (396)	0.22 (247)	$p=0.3349$					

Table 84 | Codon Usage from Domains and Non-Domains for the Excavate *G. lamblia*

Amino Acids	Codons	Domain	Non-Domain	Chi-Sqaured
Phe	UUU	0.27 (205)	0.33 (73)	p=0.1410
	UUC	0.73 (544)	0.67 (150)	p=0.1410
Leu	UUA	0.13 (163)	0.09 (39)	p=0.0480
	UUG	0.07 (90)	0.10 (41)	p=0.1182
	CUU	0.24 (303)	0.26 (112)	p=0.3741
	CUC	0.44 (551)	0.35 (147)	p=0.0011
	CUA	0.03 (37)	0.09 (37)	p<0.0001
	CUG	0.09 (113)	0.11 (48)	p=0.1885
Val	GUU	0.38 (480)	0.37 (115)	p=0.8398
	GUC	0.34 (430)	0.27 (84)	p=0.0249
	GUA	0.14 (180)	0.22 (70)	p=0.0004
	GUG	0.14 (184)	0.14 (43)	p=0.8349
Pro	CCU	0.33 (196)	0.32 (84)	p=0.7309
	CCC*	0.46 (270)	0.31 (81)	p<0.0001
	CCA	0.17 (101)	0.24 (63)	p=0.0276
	CCG	0.04 (24)	0.14 (37)	p<0.0001
His	CAU	0.30 (96)	0.35 (38)	p=0.3769
	CAC	0.70 (227)	0.65 (71)	p=0.3769
Glu	GAA	0.46 (531)	0.56 (217)	p=0.0021
	GAG	0.54 (612)	0.44 (173)	p=0.0021
Asn	AAU	0.29 (248)	0.37 (85)	p=0.0328
	AAC	0.71 (605)	0.63 (147)	p=0.0328
Lys	AAA	0.28 (410)	0.32 (146)	p=0.0866
	AAG	0.72 (1049)	0.68 (304)	p=0.0866
Cys	UGU	0.22 (116)	0.30 (49)	p=0.0485
	UGC	0.78 (402)	0.70 (112)	p=0.0485
Arg	CGU*	0.33 (263)	0.20 (63)	p<0.0001
	CGC	0.27 (213)	0.21 (66)	p=0.0451
	CGA	0.01 (7)	0.03 (8)	p=0.0636
	CGG	0.03 (25)	0.07 (23)	p=0.0040
	AGA	0.27 (217)	0.34 (106)	p=0.0501
	AGG	0.08 (67)	0.16 (50)	p=0.0005

Amino Acids	Codons	Domain	Non-Domain	Chi-Sqaured
Ser	UCU	0.32 (335)	0.25 (92)	p=0.0136
	UCC	0.36 (379)	0.30 (110)	p=0.0319
	UCA	0.09 (92)	0.12 (45)	p=0.0713
	UCG	0.06 (63)	0.09 (32)	p=0.1023
	AGU	0.04 (47)	0.07 (27)	p=0.0495
	AGC	0.13 (140)	0.18 (66)	p=0.0422
Thr	ACU	0.33 (342)	0.29 (90)	p=0.2457
	ACC*	0.50 (524)	0.37 (113)	p<0.0001
	ACA	0.09 (90)	0.18 (55)	p<0.0001
	ACG	0.08 (82)	0.16 (50)	p<0.0001
Tyr	UAU	0.44 (229)	0.48 (81)	p=0.4047
	UAC	0.56 (293)	0.52 (88)	p=0.4047
Ala	GCU	0.36 (530)	0.36 (193)	p=0.9370
	GCC	0.43 (642)	0.39 (211)	p=0.0724
	GCA	0.11 (163)	0.14 (75)	p=0.1003
	GCG	0.10 (143)	0.12 (64)	p=0.1920
Ile	AUU	0.29 (368)	0.39 (109)	p=0.0039
	AUC*	0.64 (794)	0.45 (128)	p<0.0001
	AUA	0.07 (86)	0.16 (46)	p<0.0001
Asp	GAU	0.43 (404)	0.40 (131)	p=0.4039
	GAC	0.57 (529)	0.60 (193)	p=0.4039
Gln	CAA	0.14 (89)	0.18 (62)	p=0.0861
	CAG	0.57 (371)	0.50 (172)	p=0.0383
	UAA	0.15 (99)	0.16 (56)	p=0.7335
	UAG	0.14 (89)	0.15 (53)	p=0.5230
Gly	GGU	0.46 (546)	0.40 (145)	p=0.0485
	GGC	0.37 (447)	0.31 (113)	p=0.0274
	GGA	0.12 (138)	0.21 (75)	p<0.0001
	GGG	0.05 (62)	0.09 (32)	p=0.0173

Table 85 | Codon Usage from Domains and Non-Domains for the Excavate *S. salmonicida*

Amino Acids	Codons	Domain	Non-Domain	Chi-Sqaured
Phe	UUU	0.21 (71)	0.33 (51)	p=0.0031
	UUC	0.79 (271)	0.67 (103)	p=0.0031
Leu	UUA	0.01 (8)	0.02 (9)	p=0.1035
	UUG	0.08 (55)	0.10 (38)	p=0.1661
	CUU	0.12 (84)	0.15 (56)	p=0.1308
	CUC	0.29 (207)	0.28 (104)	p=0.6937
	CUA	0.03 (18)	0.05 (19)	p=0.0266
	CUG	0.48 (347)	0.40 (150)	p=0.0083
Val	GUU	0.13 (94)	0.14 (47)	p=0.6705
	GUC	0.26 (180)	0.30 (98)	p<0.0001
	GUA	0.04 (25)	0.07 (23)	p=0.0140
	GUG	0.58 (405)	0.49 (160)	p=0.0086
Pro	CCU	0.11 (38)	0.16 (45)	p=0.0649
	CCC	0.26 (88)	0.26 (71)	p=0.9567
	CCA	0.08 (29)	0.15 (41)	p=0.0142
	CCG	0.55 (187)	0.44 (121)	p=0.0057
His	CAU	0.13 (24)	0.19 (36)	p=0.1312
	CAC	0.87 (154)	0.81 (150)	p=0.1312
Gln	CAA	0.07 (18)	0.13 (24)	p=0.0166
	CAG	0.93 (250)	0.87 (154)	p=0.0166
Asn	AAU	0.14 (38)	0.14 (20)	p=0.8196
	AAC	0.86 (242)	0.86 (119)	p=0.8196
Lys	AAA	0.06 (34)	0.09 (32)	p=0.0610
	AAG	0.94 (529)	0.91 (309)	p=0.0610
Cys	UGU	0.13 (19)	0.30 (42)	p=0.0002
	UGC*	0.87 (132)	0.70 (97)	p=0.0002
Arg	CGU	0.17 (89)	0.22 (91)	p=0.0485
	CGC*	0.71 (382)	0.50 (208)	p<0.0001
	CGA	0.03 (16)	0.07 (29)	p=0.0046
	CGG	0.06 (31)	0.11 (45)	p=0.0053
	AGA	0.01 (4)	0.06 (24)	p<0.0001
	AGG	0.02 (13)	0.05 (23)	p=0.0142
Ser	UCU	0.11 (62)	0.18 (68)	p=0.0016
	UCC	0.22 (125)	0.23 (87)	p=0.6621
	UCA	0.04 (22)	0.09 (35)	p=0.0006
	UCG	0.30 (168)	0.21 (79)	p=0.0037
	AGU	0.05 (26)	0.06 (24)	p=0.2203
	AGC	0.29 (164)	0.22 (81)	p=0.0129
Thr	ACU	0.14 (68)	0.14 (41)	p=0.9352
	ACC	0.32 (160)	0.28 (83)	p=0.2294
	ACA	0.09 (46)	0.25 (75)	p<0.0001
	ACG*	0.45 (226)	0.33 (98)	p=0.0007
Tyr	UAU	0.07 (21)	0.24 (34)	p<0.0001
	UAC*	0.93 (261)	0.76 (110)	p<0.0001
Ala	GCU	0.17 (113)	0.22 (124)	p=0.0513
	GCC*	0.42 (272)	0.30 (169)	p<0.0001
	GCA	0.11 (70)	0.13 (74)	p=0.2250
	GCG	0.30 (198)	0.36 (204)	p=0.0446
Ile	AUU	0.26 (109)	0.27 (45)	p=0.6573
	AUC*	0.71 (301)	0.55 (90)	p=0.0002
	AUA	0.04 (15)	0.18 (29)	p<0.0001
Asp	GAU	0.26 (93)	0.30 (65)	p=0.2903
	GAC	0.74 (261)	0.70 (149)	p=0.2903
Glu	GAA	0.13 (58)	0.17 (54)	p=0.2163
	GAG	0.87 (375)	0.83 (271)	p=0.2163
Gly	GGU	0.26 (146)	0.27 (80)	p=0.7720
	GGC	0.60 (337)	0.51 (152)	p=0.0126
	GGA	0.06 (36)	0.08 (24)	p=0.3630
	GGG	0.08 (44)	0.14 (42)	p=0.0035

Table 86 | Codon Usage from Domains and Non-Domains for the Excavate *L. major*

Amino Acids	Codons	Domain	Non-Domain	Chi-Sqaured
Phe	UUU	0.10 (24)	0.34 (28)	p<0.0001
	UUC*	0.90 (207)	0.66 (55)	p<0.0001
Leu	UUA	0.00 (1)	0.02 (3)	p=0.0116
	UUG	0.02 (11)	0.14 (21)	p<0.0001
	CUU	0.07 (35)	0.14 (21)	p=0.0049
	CUC	0.22 (111)	0.24 (36)	p=0.5104
	CUA	0.02 (8)	0.04 (6)	p=0.0651
	CUG*	0.67 (344)	0.41 (61)	p<0.0001
Val	GUU	0.08 (41)	0.11 (17)	p=0.2087
	GUC	0.25 (129)	0.29 (44)	p=0.3195
	GUA	0.01 (5)	0.08 (13)	p<0.0001
	GUG*	0.66 (346)	0.52 (79)	p=0.0009
Pro	CCU	0.13 (35)	0.18 (34)	p=0.1255
	CCC	0.26 (71)	0.28 (53)	p=0.6286
	CCA	0.03 (8)	0.10 (19)	p=0.0013
	CCG	0.58 (156)	0.43 (81)	p=0.0023
His	CAU	0.04 (6)	0.15 (7)	p=0.0159
	CAC	0.96 (133)	0.85 (41)	p=0.0159
Gln	CAA	0.09 (18)	0.22 (22)	p=0.0008
	CAG*	0.91 (190)	0.78 (76)	p=0.0008
Asn	AAU	0.07 (20)	0.28 (21)	p<0.0001
	AAC*	0.93 (257)	0.72 (53)	p<0.0001
Lys	AAA	0.02 (9)	0.07 (14)	p=0.0027
	AAG	0.98 (418)	0.93 (189)	p=0.0027
Cys	UGU	0.05 (8)	0.17 (9)	p=0.0070
	UGC	0.95 (143)	0.83 (43)	p=0.0070
Arg	CGU	0.14 (56)	0.25 (56)	p=0.0004
	CGC*	0.78 (320)	0.45 (101)	p<0.0001
	CGA	0.01 (6)	0.05 (11)	p=0.0105
	CGG	0.04 (17)	0.12 (26)	p=0.0004
	AGA	0.01 (3)	0.04 (8)	p=0.0091
	AGG	0.02 (10)	0.11 (24)	p<0.0001

Amino Acids	Codons	Domain	Non-Domain	Chi-Sqaured
Ser	UCU	0.09 (32)	0.13 (25)	p=0.1474
	UCC	0.24 (85)	0.24 (46)	p=0.9822
	UCA	0.02 (7)	0.06 (11)	p=0.0192
	UCG	0.30 (106)	0.20 (37)	p=0.0064
	AGU	0.02 (7)	0.07 (13)	p=0.0044
	AGC	0.32 (111)	0.30 (57)	p=0.6783
Thr	ACU	0.07 (27)	0.11 (13)	p=0.1820
	ACC	0.32 (119)	0.35 (41)	p=0.5275
	ACA	0.03 (12)	0.06 (7)	p=0.1764
	ACG	0.58 (215)	0.48 (56)	p=0.0635
Tyr	UAU	0.05 (11)	0.23 (12)	p<0.0001
	UAC*	0.95 (197)	0.77 (41)	p<0.0001
Ala	GCU	0.12 (66)	0.22 (85)	p<0.0001
	GCC	0.39 (210)	0.38 (146)	p=0.8442
	GCA	0.03 (16)	0.07 (28)	p=0.0022
	GCG*	0.46 (247)	0.32 (122)	p<0.0001
Ile	AUU	0.18 (58)	0.17 (12)	p=0.7568
	AUC	0.81 (255)	0.68 (48)	p=0.0115
	AUA	0.00 (1)	0.15 (11)	p<0.0001
Asp	GAU	0.18 (47)	0.27 (34)	p=0.0376
	GAC	0.82 (212)	0.73 (90)	p=0.0376
Glu	GAA	0.08 (23)	0.18 (31)	p=0.0011
	GAG	0.92 (262)	0.82 (138)	p=0.0011
Gly	GGU	0.27 (128)	0.35 (84)	p=0.0207
	GGC*	0.64 (309)	0.43 (104)	p<0.0001
	GGA	0.04 (18)	0.10 (24)	p=0.0007
	GGG	0.05 (25)	0.12 (28)	p=0.0018

Table 87 | Codon Usage from Domains and Non-Domains for the Excavate *L. pyrrhocoris*

Amino Acids	Codons	Domain	Non-Domain	Chi-Sqaured	Amino Acids	Codons	Domain	Non-Domain	Chi-Sqaured
Phe	UUU	0.43 (347)	0.47 (461)	p=0.0918	**Ser**	UCU	0.13 (245)	0.15 (643)	p=0.0695
	UUC	0.57 (465)	0.53 (526)	p=0.0918		UCC	0.19 (344)	0.20 (841)	p=0.3394
Leu	UUA	0.03 (85)	0.04 (142)	p=0.0639		UCA	0.07 (133)	0.10 (440)	p<0.0001
	UUG	0.14 (356)	0.16 (515)	p=0.0996		UCG	0.21 (390)	0.20 (851)	p=0.2796
	CUU	0.12 (303)	0.13 (408)	p=0.6447		AGU	0.11 (192)	0.09 (382)	p=0.0720
	CUC	0.24 (601)	0.22 (698)	p=0.0184		AGC	0.28 (512)	0.25 (1046)	p=0.0072
	CUA	0.05 (118)	0.06 (200)	p=0.0195	**Thr**	ACU	0.10 (146)	0.15 (365)	p<0.0001
	CUG	0.41 (1014)	0.39 (1265)	p=0.1815		ACC	0.31 (442)	0.28 (672)	p=0.0156
Val	GUU	0.12 (230)	0.14 (369)	p=0.0599		ACA	0.16 (231)	0.19 (452)	p=0.0813
	GUC	0.23 (420)	0.23 (586)	p=0.8985		ACG	0.42 (588)	0.39 (934)	p=0.0480
	GUA	0.08 (153)	0.10 (253)	p=0.0693	**Tyr**	UAU	0.18 (125)	0.22 (183)	p=0.0530
	GUG	0.57 (1045)	0.53 (1352)	p=0.0140		UAC	0.82 (571)	0.78 (651)	p=0.0530
Pro	CCU	0.16 (189)	0.19 (531)	p=0.0215	**Ala**	GCU	0.14 (342)	0.18 (833)	p<0.0001
	CCC	0.23 (262)	0.24 (661)	p=0.2866		GCC*	0.31 (768)	0.27 (1262)	p=0.0007
	CCA	0.19 (224)	0.21 (577)	p=0.2059		GCA	0.17 (430)	0.21 (957)	p=0.0009
	CCG*	0.42 (487)	0.35 (970)	p<0.0001		GCG	0.37 (919)	0.34 (1553)	p=0.0022
His	CAU	0.23 (158)	0.30 (307)	p=0.0020	**Ile**	AUU	0.38 (363)	0.33 (355)	p=0.0179
	CAC	0.77 (518)	0.70 (709)	p=0.0020		AUC	0.56 (538)	0.56 (605)	p=0.9451
Gln	CAA	0.20 (199)	0.25 (437)	p=0.0068		AUA	0.07 (63)	0.12 (127)	p<0.0001
	CAG	0.80 (784)	0.75 (1326)	p=0.0068	**Asp**	GAU	0.39 (477)	0.39 (742)	p=0.9970
Asn	AAU	0.26 (183)	0.28 (340)	p=0.4095		GAC	0.61 (750)	0.61 (1167)	p=0.9970
	AAC	0.74 (511)	0.72 (869)	p=0.4095	**Glu**	GAA	0.21 (337)	0.26 (641)	p=0.0004
Lys	AAA	0.18 (169)	0.19 (248)	p=0.2711		GAG*	0.79 (1250)	0.74 (1813)	p=0.0004
	AAG	0.82 (791)	0.81 (1028)	p=0.2711	**Gly**	GGU	0.21 (302)	0.21 (501)	p=0.9088
Cys	UGU	0.20 (101)	0.27 (160)	p=0.0050		GGC*	0.51 (745)	0.45 (1080)	p=0.0002
	UGC*	0.80 (407)	0.73 (430)	p=0.0050		GGA	0.10 (146)	0.14 (338)	p=0.0002
Arg	CGU	0.19 (329)	0.19 (528)	p=0.8993		GGG	0.19 (273)	0.21 (495)	p=0.1534
	CGC*	0.44 (765)	0.38 (1039)	p<0.0001					
	CGA	0.10 (174)	0.12 (316)	p=0.1391					
	CGG	0.19 (326)	0.18 (497)	p=0.5005					
	AGA	0.02 (39)	0.05 (143)	p<0.0001					
	AGG	0.05 (88)	0.08 (217)	p=0.0003					

Table 88 | Codon Usage from Domains and Non-Domains for the Excavate *L. seymouri*

Amino Acids	Codons	Domain	Non-Domain	Chi-Sqaured	Amino Acids	Codons	Domain	Non-Domain	Chi-Sqaured
Phe	UUU	0.35 (149)	0.56 (94)	p<0.0001	**Ser**	UCU	0.17 (92)	0.18 (61)	p=0.7681
	UUC*	0.65 (281)	0.44 (73)	p<0.0001		UCC	0.22 (121)	0.16 (53)	p=0.0192
Leu	UUA	0.03 (26)	0.10 (37)	p<0.0001		UCA	0.14 (74)	0.13 (45)	p=0.9515
	UUG	0.17 (141)	0.19 (69)	p=0.3560		UCG	0.17 (90)	0.16 (53)	p=0.7730
	CUU	0.27 (228)	0.22 (81)	p=0.1102		AGU	0.12 (65)	0.15 (52)	p=0.1870
	CUC	0.25 (211)	0.19 (69)	p=0.0320		AGC	0.19 (103)	0.23 (77)	p=0.2152
	CUA	0.08 (69)	0.09 (31)	p=0.9033	**Thr**	ACU	0.21 (132)	0.26 (65)	p=0.1652
	CUG	0.20 (171)	0.21 (75)	p=0.9033		ACC	0.25 (155)	0.21 (54)	p=0.3190
Val	GUU	0.26 (220)	0.31 (86)	p=0.1308		ACA	0.24 (147)	0.32 (81)	p=0.0116
	GUC	0.19 (164)	0.19 (53)	p=0.9690		ACG*	0.30 (189)	0.21 (52)	p=0.0047
	GUA	0.13 (107)	0.13 (35)	p=0.9743	**Tyr**	UAU	0.31 (92)	0.43 (38)	p=0.0553
	GUG	0.42 (360)	0.38 (106)	p=0.2146		UAC	0.69 (201)	0.57 (50)	p=0.0553
Pro	CCU	0.18 (59)	0.25 (70)	p=0.0296	**Ala**	GCU	0.24 (218)	0.29 (161)	p=0.0863
	CCC	0.32 (107)	0.34 (93)	p=0.7710		GCC	0.29 (260)	0.33 (184)	p=0.1657
	CCA	0.22 (74)	0.22 (60)	p=0.9453		GCA	0.24 (212)	0.22 (125)	p=0.5426
	CCG	0.28 (93)	0.19 (54)	p=0.0198		GCG	0.23 (205)	0.17 (94)	p=0.0050
His	CAU	0.32 (76)	0.46 (46)	p=0.0162	**Ile**	AUU	0.47 (264)	0.46 (76)	p=0.8909
	CAC	0.68 (162)	0.54 (53)	p=0.0162		AUC	0.43 (244)	0.39 (64)	p=0.3244
Gln	CAA	0.32 (109)	0.51 (112)	p<0.0001		AUA	0.09 (53)	0.15 (25)	p=0.0528
	CAG*	0.68 (229)	0.49 (109)	p<0.0001	**Asp**	GAU	0.48 (244)	0.53 (133)	p=0.2369
Asn	AAU	0.36 (152)	0.43 (70)	p=0.1089		GAC	0.52 (263)	0.47 (118)	p=0.2369
	AAC	0.64 (272)	0.57 (91)	p=0.1089	**Glu**	GAA	0.39 (271)	0.44 (155)	p=0.1588
Lys	AAA	0.26 (209)	0.27 (105)	p=0.8548		GAG	0.61 (425)	0.56 (200)	p=0.1588
	AAG	0.74 (600)	0.73 (291)	p=0.8548	**Gly**	GGU	0.42 (314)	0.34 (111)	p=0.0108
Cys	UGU	0.32 (46)	0.41 (28)	p=0.3100		GGC	0.32 (236)	0.30 (98)	p=0.5739
	UGC	0.68 (96)	0.59 (41)	p=0.3100		GGA	0.14 (104)	0.21 (69)	p=0.0054
Arg	CGU	0.30 (236)	0.28 (112)	p=0.4948		GGG	0.12 (92)	0.16 (52)	p=0.1543
	CGC	0.42 (329)	0.35 (139)	p=0.0195					
	CGA	0.07 (56)	0.07 (28)	p=0.9280					
	CGG	0.11 (88)	0.11 (46)	p=0.9642					
	AGA	0.03 (22)	0.07 (28)	p=0.0012					
	AGG	0.07 (55)	0.12 (48)	p=0.0056					

Table 89 | Codon Usage from Domains and Non-Domains for the Excavate *T. brucei*

Amino Acids	Codons	Domain	Non-Domain	Chi-Sqaured
Phe	UUU	0.56 (295)	0.73 (187)	p<0.0001
	UUC*	0.44 (231)	0.27 (69)	p<0.0001
Leu	UUA	0.04 (51)	0.06 (33)	p=0.1695
	UUG	0.16 (201)	0.21 (120)	p=0.0231
	CUU	0.23 (283)	0.18 (105)	p=0.0255
	CUC	0.19 (239)	0.20 (116)	p=0.7644
	CUA	0.03 (38)	0.03 (20)	p=0.7713
	CUG	0.34 (415)	0.31 (180)	p=0.3261
Val	GUU	0.23 (291)	0.27 (129)	p=0.0738
	GUC	0.20 (254)	0.18 (88)	p=0.5387
	GUA	0.07 (86)	0.07 (34)	p=0.8562
	GUG	0.51 (656)	0.48 (230)	p=0.2597
Pro	CCU	0.19 (149)	0.21 (83)	p=0.3321
	CCC	0.25 (200)	0.27 (104)	p=0.6154
	CCA	0.23 (181)	0.25 (96)	p=0.5178
	CCG	0.33 (260)	0.27 (104)	p=0.0415
His	CAU	0.32 (100)	0.47 (68)	p=0.0027
	CAC	0.68 (216)	0.53 (78)	p=0.0027
Gln	CAA	0.35 (187)	0.44 (89)	p=0.0298
	CAG	0.65 (344)	0.56 (112)	p=0.0298
Asn	AAU	0.45 (329)	0.50 (100)	p=0.2438
	AAC	0.55 (409)	0.50 (102)	p=0.2438
Lys	AAA	0.33 (344)	0.31 (117)	p=0.6426
	AAG	0.67 (714)	0.69 (260)	p=0.6426
Cys	UGU	0.23 (63)	0.51 (108)	p<0.0001
	UGC*	0.77 (206)	0.49 (105)	p<0.0001
Arg	CGU	0.23 (190)	0.22 (84)	p=0.8186
	CGC*	0.34 (283)	0.21 (80)	p<0.0001
	CGA	0.05 (41)	0.09 (34)	p=0.0106
	CGG	0.10 (84)	0.14 (52)	p=0.0848
	AGA	0.08 (67)	0.11 (40)	p=0.1983
	AGG	0.20 (165)	0.24 (90)	p=0.1527
Ser	UCU	0.17 (223)	0.25 (122)	p=0.0001
	UCC	0.18 (234)	0.20 (96)	p=0.3534
	UCA	0.13 (168)	0.13 (63)	p=0.9512
	UCG*	0.17 (222)	0.09 (43)	p<0.0001
	AGU	0.16 (211)	0.15 (72)	p=0.5999
	AGC	0.21 (277)	0.19 (95)	p=0.5529
Thr	ACU	0.19 (227)	0.21 (113)	p=0.5623
	ACC	0.21 (252)	0.26 (143)	p=0.0367
	ACA	0.22 (265)	0.23 (126)	p=0.8769
	ACG	0.37 (439)	0.31 (169)	p=0.0104
Tyr	UAU	0.24 (105)	0.44 (41)	p<0.0001
	UAC*	0.76 (338)	0.56 (52)	p=0.0001
Ala	GCU	0.19 (274)	0.27 (205)	p<0.0001
	GCC	0.29 (415)	0.26 (200)	p=0.1991
	GCA	0.21 (293)	0.21 (158)	p=0.9078
	GCG	0.31 (445)	0.26 (195)	p=0.0088
Ile	AUU	0.44 (267)	0.48 (87)	p=0.3963
	AUC*	0.47 (286)	0.29 (53)	p<0.0001
	AUA	0.09 (52)	0.23 (41)	p<0.0001
Asp	GAU	0.44 (376)	0.56 (128)	p=0.0018
	GAC	0.56 (470)	0.44 (99)	p=0.0018
Glu	GAA	0.38 (449)	0.42 (169)	p=0.1600
	GAG	0.62 (726)	0.58 (230)	p=0.1600
Gly	GGU	0.26 (373)	0.27 (124)	p=0.7346
	GGC*	0.38 (542)	0.25 (114)	p<0.0001
	GGA	0.20 (282)	0.24 (112)	p=0.0370
	GGG	0.17 (240)	0.24 (111)	p=0.0005

Table 90 | Codon Usage from Domains and Non-Domains for the Excavate *T. cruzi*

Amino Acids	Codons	Domain	Non-Domain	Chi-Sqaured	Amino Acids	Codons	Domain	Non-Domain	Chi-Sqaured
Phe	UUU	0.43 (149)	0.57 (50)	p=0.0284	**Ser**	UCU	0.19 (105)	0.24 (63)	p=0.1499
	UUC	0.57 (197)	0.43 (38)	p=0.0284		UCC	0.16 (87)	0.17 (46)	p=0.6574
Leu	UUA	0.05 (36)	0.07 (23)	p=0.2118		UCA	0.11 (63)	0.13 (35)	p=0.5518
	UUG	0.15 (116)	0.19 (65)	p=0.1260		UCG	0.13 (70)	0.07 (18)	p=0.0145
	CUU	0.19 (144)	0.14 (48)	p=0.0659		AGU	0.16 (91)	0.18 (47)	p=0.7551
	CUC	0.25 (190)	0.21 (73)	p=0.2458		AGC	0.25 (137)	0.22 (58)	p=0.3820
	CUA	0.04 (31)	0.04 (13)	p=0.9825	**Thr**	ACU	0.19 (121)	0.28 (65)	p=0.0053
	CUG	0.32 (248)	0.35 (119)	p=0.4596		ACC	0.24 (153)	0.21 (49)	p=0.4278
Val	GUU	0.22 (162)	0.25 (74)	p=0.4273		ACA	0.33 (213)	0.30 (70)	p=0.4163
	GUC	0.16 (115)	0.17 (51)	p=0.6783		ACG	0.24 (151)	0.21 (48)	p=0.4045
	GUA	0.11 (80)	0.11 (32)	p=0.9873	**Tyr**	UAU	0.31 (90)	0.41 (28)	p=0.1581
	GUG	0.51 (377)	0.48 (144)	p=0.3367		UAC	0.69 (202)	0.59 (41)	p=0.1581
Pro	CCU	0.19 (70)	0.31 (61)	p=0.0015	**Ala**	GCU	0.24 (239)	0.31 (110)	p=0.0258
	CCC	0.30 (111)	0.22 (42)	p=0.0363		GCC	0.31 (301)	0.32 (114)	p=0.7742
	CCA	0.31 (115)	0.28 (55)	p=0.5123		GCA	0.25 (246)	0.18 (65)	p=0.0088
	CCG	0.19 (70)	0.19 (36)	p=0.9600		GCG	0.20 (193)	0.19 (70)	p=0.9919
His	CAU	0.39 (91)	0.36 (44)	p=0.6847	**Ile**	AUU	0.48 (223)	0.52 (61)	p=0.5471
	CAC	0.61 (143)	0.64 (78)	p=0.6847		AUC	0.42 (195)	0.33 (39)	p=0.0949
Gln	CAA	0.26 (101)	0.43 (59)	p=0.0003		AUA	0.10 (46)	0.15 (18)	p=0.1360
	CAG*	0.74 (283)	0.57 (77)	p=0.0003	**Asp**	GAU	0.50 (261)	0.49 (76)	p=0.9602
Asn	AAU	0.40 (177)	0.46 (77)	p=0.2189		GAC	0.50 (261)	0.51 (78)	p=0.9602
	AAC	0.60 (269)	0.54 (92)	p=0.2189	**Glu**	GAA	0.40 (360)	0.50 (116)	p=0.0132
Lys	AAA	0.30 (275)	0.25 (66)	p=0.2039		GAG	0.60 (529)	0.50 (117)	p=0.0132
	AAG	0.70 (652)	0.75 (194)	p=0.2039	**Gly**	GGU	0.36 (218)	0.38 (92)	p=0.5148
Cys	UGU	0.32 (74)	0.53 (48)	p=0.0009		GGC	0.31 (189)	0.32 (77)	p=0.8026
	UGC*	0.68 (154)	0.47 (42)	p=0.0009		GGA	0.20 (124)	0.15 (37)	p=0.1254
Arg	CGU	0.23 (155)	0.28 (56)	p=0.1520		GGG	0.14 (83)	0.15 (35)	p=0.7848
	CGC	0.40 (269)	0.33 (65)	p=0.0856					
	CGA	0.04 (27)	0.09 (18)	p=0.0078					
	CGG	0.08 (57)	0.08 (15)	p=0.8032					
	AGA	0.11 (71)	0.12 (24)	p=0.6168					
	AGG	0.14 (95)	0.10 (20)	p=0.1800					

Table 91 | Codon Usage from Domains and Non-Domains for the Excavate *T. theileri*

Amino Acids	Codons	Domain	Non-Domain	Chi-Sqaured	Amino Acids	Codons	Domain	Non-Domain	Chi-Sqaured
Phe	UUU	0.18 (134)	0.32 (83)	p<0.0001	**Ser**	UCU*	0.49 (480)	0.37 (165)	p<0.0001
	UUC*	0.82 (591)	0.68 (177)	p<0.0001		UCC*	0.23 (226)	0.15 (67)	p=0.0004
Leu	UUA	0.29 (443)	0.36 (185)	p=0.0023		UCA	0.09 (88)	0.14 (63)	p=0.0038
	UUG	0.20 (313)	0.16 (85)	p=0.0577		UCG	0.00 (3)	0.00 (2)	p=0.6773
	CUU	0.24 (370)	0.24 (123)	p=0.9479		AGU	0.09 (86)	0.15 (66)	p=0.0007
	CUC	0.25 (393)	0.20 (105)	p=0.0187		AGC	0.09 (88)	0.18 (81)	p<0.0001
	CUA	0.01 (23)	0.03 (17)	p=0.0102	**Thr**	ACU	0.52 (515)	0.53 (194)	p=0.9798
	CUG	0.00 (4)	0.00 (2)	p=0.6396		ACC	0.39 (385)	0.37 (137)	p=0.4763
Val	GUU	0.53 (630)	0.54 (206)	p=0.6401		ACA	0.08 (80)	0.10 (38)	p=0.2140
	GUC	0.40 (480)	0.32 (120)	p=0.0024		ACG	0.00 (1)	0.00 (0)	p=0.5395
	GUA	0.05 (58)	0.11 (42)	p<0.0001	**Tyr**	UAU	0.39 (243)	0.37 (65)	p=0.7039
	GUG	0.02 (18)	0.03 (10)	p=0.1499		UAC	0.61 (381)	0.63 (109)	p=0.7039
Pro	CCU	0.54 (331)	0.56 (117)	p=0.6825	**Ala**	GCU	0.73 (1051)	0.71 (373)	p=0.6376
	CCC	0.39 (241)	0.37 (78)	p=0.5661		GCC	0.23 (330)	0.22 (114)	p=0.6610
	CCA	0.06 (39)	0.07 (14)	p=0.8810		GCA	0.05 (68)	0.06 (33)	p=0.1478
	CCG	0.00 (1)	0.00 (1)	p=0.4273		GCG	0.00 (0)	0.00 (2)	p=0.0184
His	CAU	0.27 (88)	0.39 (38)	p=0.0209	**Ile**	AUU	0.52 (615)	0.51 (207)	p=0.6628
	CAC	0.73 (242)	0.61 (60)	p=0.0209		AUC	0.43 (511)	0.35 (142)	p=0.0030
Glu	GAA*	0.98 (1010)	0.93 (371)	p<0.0001		AUA	0.05 (57)	0.14 (59)	p<0.0001
	GAG	0.02 (25)	0.07 (28)	p<0.0001	**Asp**	GAU	0.65 (572)	0.67 (206)	p=0.5988
Asn	AAU	0.33 (312)	0.44 (265)	p<0.0001		GAC	0.35 (302)	0.33 (101)	p=0.5988
	AAC*	0.67 (628)	0.56 (335)	p<0.0001	**Glu**	CAA	0.46 (339)	0.41 (206)	p=0.0423
Lys	AAA	0.21 (291)	0.38 (260)	p<0.0001		CAG	0.00 (4)	0.01 (5)	p=0.2480
	AAG*	0.79 (1099)	0.62 (430)	p<0.0001		UAA	0.46 (396)	0.49 (248)	p=0.2825
Cys	UGU	0.34 (68)	0.37 (20)	p=0.6949		UAG	0.07 (60)	0.09 (46)	p=0.1571
	UGC	0.66 (131)	0.63 (34)	p=0.6949	**Gly**	GGU*	0.84 (1058)	0.65 (220)	p<0.0001
Arg	CGU	0.09 (67)	0.07 (19)	p=0.3330		GGC	0.05 (67)	0.12 (41)	p<0.0001
	CGC	0.00 (3)	0.00 (0)	p=0.2975		GGA	0.10 (129)	0.22 (73)	p<0.0001
	CGA	0.00 (0)	0.00 (1)	p=0.0958		GGG	0.00 (3)	0.01 (5)	p=0.0042
	CGG	0.00 (0)	0.00 (0)	p=0.0963					
	AGA	0.91 (707)	0.88 (247)	p=0.1711					
	AGG	0.00 (2)	0.05 (14)	p<0.0001					

Table 92 | Codon Usage from Domains and Non-Domains for the Alveolate *T. thermophila*

Amino Acids	Codons	Domain	Non-Domain	Chi-Sqaured
Phe	UUU	0.52 (358)	0.59 (102)	p=0.1115
	UUC	0.48 (332)	0.41 (72)	p=0.1115
Leu	UUA	0.24 (324)	0.27 (106)	p=0.1881
	UUG	0.17 (227)	0.16 (62)	p=0.6866
	CUU	0.37 (497)	0.33 (129)	p=0.1892
	CUC	0.09 (119)	0.09 (34)	p=0.9667
	CUA	0.10 (134)	0.09 (35)	p=0.5888
	CUG	0.04 (50)	0.06 (23)	p=0.0552
Val	GUU	0.50 (584)	0.48 (215)	p=0.6395
	GUC	0.21 (242)	0.16 (72)	p=0.0463
	GUA	0.18 (213)	0.24 (109)	p=0.0040
	GUG	0.12 (138)	0.11 (49)	p=0.6881
Pro	CCU	0.46 (274)	0.46 (113)	p=0.8937
	CCC	0.04 (25)	0.04 (10)	p=0.9098
	CCA	0.47 (280)	0.46 (113)	p=0.6877
	CCG	0.02 (11)	0.04 (10)	p=0.0639
His	CAU	0.69 (183)	0.75 (60)	p=0.3283
	CAC	0.31 (81)	0.25 (20)	p=0.3283
Gln	CAA	0.80 (445)	0.77 (167)	p=0.2323
	CAG	0.20 (108)	0.23 (51)	p=0.2323
Asn	AAU	0.60 (454)	0.70 (178)	p=0.0027
	AAC	0.40 (305)	0.30 (75)	p=0.0027
Lys	AAA	0.65 (728)	0.64 (311)	p=0.6361
	AAG	0.35 (395)	0.36 (178)	p=0.6361
Cys	UGU	0.58 (145)	0.36 (14)	p=0.0115
	UGC	0.42 (107)	0.64 (25)	p=0.0115
Arg	CGU	0.14 (101)	0.12 (36)	p=0.4646
	CGC	0.04 (31)	0.03 (10)	p=0.5145
	CGA	0.03 (23)	0.04 (12)	p=0.4753
	CGG	0.00 (2)	0.02 (5)	p=0.0128
	AGA	0.67 (484)	0.65 (189)	p=0.4122
	AGG	0.11 (77)	0.14 (40)	p=0.1806
Ser	UCU	0.21 (212)	0.21 (77)	p=0.8568
	UCC	0.06 (59)	0.08 (29)	p=0.1382
	UCA	0.35 (359)	0.28 (103)	p=0.0185
	UCG	0.02 (23)	0.05 (18)	p=0.0090
	AGU	0.22 (228)	0.23 (85)	p=0.6713
	AGC	0.14 (140)	0.14 (51)	p=0.8728
Thr	ACU	0.53 (525)	0.55 (247)	p=0.3474
	ACC	0.18 (177)	0.15 (69)	p=0.2831
	ACA	0.28 (280)	0.28 (123)	p=0.8346
	ACG	0.01 (14)	0.02 (7)	p=0.8102
Tyr	UAU	0.68 (327)	0.68 (79)	p=0.9235
	UAC	0.32 (154)	0.32 (38)	p=0.9235
Ala	GCU	0.62 (761)	0.62 (392)	p=0.8962
	GCC	0.12 (152)	0.11 (69)	p=0.3730
	GCA	0.24 (301)	0.25 (160)	p=0.6677
	GCG	0.02 (23)	0.02 (13)	p=0.7757
Ile	AUU	0.59 (609)	0.56 (159)	p=0.3332
	AUC	0.27 (284)	0.22 (63)	p=0.0669
	AUA	0.14 (143)	0.22 (64)	p=0.0004
Asp	GAU	0.72 (817)	0.77 (247)	p=0.0766
	GAC	0.28 (318)	0.23 (74)	p=0.0766
Glu	GAA	0.85 (1015)	0.82 (977)	p=0.0251
	GAG	0.15 (175)	0.18 (216)	p=0.0251
Gly	GGU	0.36 (424)	0.32 (134)	p=0.0956
	GGC	0.13 (156)	0.14 (57)	p=0.9331
	GGA	0.41 (474)	0.44 (187)	p=0.1785
	GGG	0.10 (115)	0.10 (44)	p=0.7294

Table 93 | Codon Usage from Domains and Non-Domains for the Alveolate *S. coeruleus*

Amino Acids	Codons	Domain	Non-Domain	Chi-Sqaured	Amino Acids	Codons	Domain	Non-Domain	Chi-Sqaured
Phe	UUU	0.90 (908)	0.84 (335)	p=0.0022	**Ser**	UCU	0.27 (384)	0.21 (167)	p=0.0050
	UUC	0.10 (99)	0.16 (62)	p=0.0022		UCC	0.05 (70)	0.04 (34)	p=0.5746
Leu	UUA	0.69 (1361)	0.68 (470)	p=0.5103		UCA	0.33 (472)	0.30 (231)	p=0.1069
	UUG	0.13 (253)	0.12 (86)	p=0.7604		UCG	0.02 (28)	0.02 (19)	p=0.4537
	CUU	0.10 (196)	0.11 (77)	p=0.3914		AGU	0.28 (407)	0.35 (273)	p=0.0014
	CUC	0.01 (16)	0.01 (4)	p=0.5362		AGC	0.05 (73)	0.07 (57)	p=0.0347
	CUA	0.07 (138)	0.07 (51)	p=0.7610	**Thr**	ACU	0.37 (397)	0.33 (151)	p=0.1425
	CUG	0.01 (12)	0.01 (8)	p=0.1536		ACC	0.06 (61)	0.09 (40)	p=0.0268
Val	GUU	0.51 (650)	0.43 (175)	p=0.0067		ACA	0.51 (557)	0.54 (248)	p=0.3946
	GUC	0.05 (58)	0.02 (9)	p=0.0373		ACG	0.06 (68)	0.05 (22)	p=0.2475
	GUA	0.37 (475)	0.46 (185)	p=0.0024	**Tyr**	UAU	0.86 (771)	0.84 (335)	p=0.3298
	GUG	0.07 (93)	0.09 (36)	p=0.2918		UAC	0.14 (121)	0.16 (62)	p=0.3298
Pro	CCU*	0.38 (259)	0.26 (105)	p<0.0001	**Ala**	GCU	0.42 (486)	0.34 (109)	p=0.0067
	CCC	0.06 (39)	0.35 (144)	p<0.0001		GCC	0.06 (63)	0.07 (22)	p=0.3550
	CCA*	0.56 (380)	0.38 (155)	p<0.0001		GCA	0.50 (570)	0.56 (180)	p=0.0421
	CCG	0.01 (5)	0.01 (6)	p=0.2409		GCG	0.02 (25)	0.03 (9)	p=0.5102
His	CAU	0.84 (357)	0.88 (123)	p=0.2358	**Ile**	AUU	0.50 (918)	0.49 (327)	p=0.6417
	CAC	0.16 (66)	0.12 (16)	p=0.2358		AUC	0.05 (91)	0.05 (31)	p=0.7455
Gln	CAA	0.88 (617)	0.89 (273)	p=0.8080		AUA	0.45 (816)	0.46 (306)	p=0.5429
	CAG	0.12 (81)	0.11 (34)	p=0.8080	**Asp**	GAU	0.90 (1203)	0.89 (660)	p=0.3863
Asn	AAU	0.85 (1321)	0.85 (993)	p=0.8189		GAC	0.10 (128)	0.11 (80)	p=0.3863
	AAC	0.15 (231)	0.15 (178)	p=0.8189	**Glu**	GAA	0.90 (1634)	0.88 (1107)	p=0.0990
Lys	AAA	0.85 (1922)	0.86 (1300)	p=0.3120		GAG	0.10 (176)	0.12 (145)	p=0.0990
	AAG	0.15 (337)	0.14 (207)	p=0.3120	**Gly**	GGU	0.35 (407)	0.35 (165)	p=0.8886
Cys	UGU	0.76 (241)	0.76 (77)	p=0.9954		GGC	0.02 (23)	0.03 (14)	p=0.2102
	UGC	0.24 (75)	0.24 (24)	p=0.9954		GGA	0.58 (673)	0.56 (259)	p=0.3830
Arg	CGU	0.09 (88)	0.06 (27)	p=0.0500		GGG	0.05 (56)	0.06 (27)	p=0.4201
	CGC	0.01 (11)	0.02 (8)	p=0.3224					
	CGA	0.04 (34)	0.03 (13)	p=0.5525					
	CGG	0.00 (2)	0.00 (1)	p=0.9491					
	AGA	0.77 (742)	0.79 (348)	p=0.6443					
	AGG	0.08 (81)	0.10 (46)	p=0.2423					

Table 94 | Codon Usage from Domains and Non-Domains for the Alveolate *P. gallinaceum*

Amino Acids	Codons	Domain	Non-Domain	Chi-Sqaured	Amino Acids	Codons	Domain	Non-Domain	Chi-Sqaured
Phe	UUU	0.85 (1215)	0.85 (1740)	p=0.8570	**Ser**	UCU	0.24 (469)	0.22 (656)	p=0.1201
	UUC	0.15 (216)	0.15 (304)	p=0.8570		UCC	0.09 (179)	0.08 (225)	p=0.0482
Leu	UUA*	0.66 (1996)	0.61 (1912)	p=0.0007		UCA	0.29 (562)	0.25 (742)	p=0.0034
	UUG	0.12 (360)	0.14 (433)	p=0.0147		UCG	0.05 (89)	0.05 (138)	p=0.8763
	CUU	0.10 (319)	0.11 (341)	p=0.5476		AGU	0.27 (521)	0.32 (945)	p<0.0001
	CUC	0.02 (62)	0.02 (74)	p=0.3637		AGC	0.06 (115)	0.08 (233)	p=0.0085
	CUA	0.08 (244)	0.09 (285)	p=0.1108	**Thr**	ACU	0.27 (378)	0.25 (428)	p=0.1127
	CUG	0.02 (59)	0.02 (64)	p=0.7416		ACC*	0.15 (209)	0.11 (188)	p=0.0005
Val	GUU	0.43 (590)	0.39 (637)	p=0.0131		ACA	0.48 (667)	0.56 (962)	p<0.0001
	GUC	0.06 (82)	0.06 (91)	p=0.5814		ACG	0.09 (128)	0.08 (145)	p=0.4078
	GUA	0.41 (560)	0.43 (703)	p=0.3323	**Tyr**	UAU	0.88 (1498)	0.88 (2215)	p=0.9232
	GUG	0.10 (131)	0.13 (210)	p=0.0061		UAC	0.12 (201)	0.12 (300)	p=0.9232
Pro	CCU	0.41 (298)	0.39 (315)	p=0.5118	**Ala**	GCU	0.44 (385)	0.41 (275)	p=0.1345
	CCC	0.07 (53)	0.12 (95)	p=0.0025		GCC	0.11 (96)	0.10 (65)	p=0.3468
	CCA	0.47 (345)	0.44 (356)	p=0.2417		GCA	0.41 (360)	0.44 (299)	p=0.3003
	CCG	0.05 (34)	0.05 (38)	p=0.9492		GCG	0.03 (27)	0.06 (39)	p=0.0108
His	CAU	0.82 (698)	0.86 (991)	p=0.0370	**Ile**	AUU	0.42 (1325)	0.38 (1579)	p=0.0044
	CAC	0.18 (151)	0.14 (166)	p=0.0370		AUC	0.07 (208)	0.07 (273)	p=0.8716
Gln	CAA	0.87 (855)	0.85 (1092)	p=0.1627		AUA	0.52 (1643)	0.55 (2257)	p=0.0067
	CAG	0.13 (128)	0.15 (194)	p=0.1627	**Asp**	GAU	0.87 (1579)	0.86 (2734)	p=0.4546
Asn	AAU	0.88 (3051)	0.86 (6745)	p=0.0285		GAC	0.13 (241)	0.14 (445)	p=0.4546
	AAC	0.12 (421)	0.14 (1065)	p=0.0285	**Glu**	GAA*	0.89 (2211)	0.85 (2954)	p<0.0001
Lys	AAA*	0.84 (2922)	0.81 (4660)	p<0.0001		GAG	0.11 (271)	0.15 (503)	p<0.0001
	AAG	0.16 (539)	0.19 (1120)	p<0.0001	**Gly**	GGU	0.39 (420)	0.40 (441)	p=0.7555
Cys	UGU	0.88 (480)	0.89 (597)	p=0.5645		GGC	0.03 (36)	0.04 (49)	p=0.1982
	UGC	0.12 (66)	0.11 (74)	p=0.5645		GGA	0.47 (505)	0.44 (479)	p=0.0706
Arg	CGU	0.11 (105)	0.10 (124)	p=0.4303		GGG	0.10 (105)	0.12 (132)	p=0.1107
	CGC	0.01 (8)	0.02 (21)	p=0.0834					
	CGA	0.08 (75)	0.08 (99)	p=0.9316					
	CGG	0.01 (11)	0.01 (8)	p=0.2007					
	AGA	0.64 (587)	0.61 (725)	p=0.1144					
	AGG	0.14 (130)	0.18 (217)	p=0.0145					

Table 95 | Codon Usage from Domains and Non-Domains for the Alveolate *P. falciparum*

Amino Acids	Codons	Domain	Non-Domain	Chi-Sqaured
Phe	UUU	0.50 (515)	0.51 (272)	p=0.8075
	UUC	0.50 (513)	0.49 (264)	p=0.8075
Leu	UUA	0.16 (321)	0.18 (215)	p=0.0896
	UUG	0.22 (439)	0.18 (213)	p=0.0126
	CUU	0.07 (146)	0.11 (137)	p<0.0001
	CUC	0.18 (375)	0.16 (193)	p=0.1143
	CUA	0.13 (272)	0.12 (141)	p=0.2164
	CUG	0.24 (487)	0.25 (293)	p=0.6499
Val	GUU	0.17 (267)	0.18 (135)	p=0.8966
	GUC	0.22 (334)	0.21 (161)	p=0.6781
	GUA	0.22 (334)	0.23 (177)	p=0.4686
	GUG	0.39 (605)	0.38 (296)	p=0.7123
Pro	CCU	0.12 (87)	0.11 (87)	p=0.9332
	CCC	0.42 (314)	0.36 (271)	p=0.0140
	CCA	0.32 (238)	0.34 (261)	p=0.2732
	CCG	0.15 (110)	0.18 (139)	p=0.0563
His	CAU	0.33 (151)	0.41 (139)	p=0.0174
	CAC	0.67 (312)	0.59 (202)	p=0.0174
Gln	CAA	0.50 (467)	0.49 (297)	p=0.5037
	CAG	0.50 (459)	0.51 (313)	p=0.5037
Asn	AAU	0.46 (731)	0.47 (573)	p=0.8073
	AAC	0.54 (850)	0.53 (654)	p=0.8073
Lys	AAA	0.50 (1340)	0.50 (987)	p=0.5375
	AAG	0.50 (1366)	0.50 (970)	p=0.5375
Cys	UGU	0.34 (141)	0.36 (107)	p=0.4912
	UGC	0.66 (275)	0.64 (187)	p=0.4912
Arg	CGU	0.05 (39)	0.07 (44)	p=0.1405
	CGC	0.09 (67)	0.14 (84)	p=0.0063
	CGA	0.12 (89)	0.11 (64)	p=0.3721
	CGG	0.06 (46)	0.07 (44)	p=0.4681
	AGA	0.37 (272)	0.30 (182)	p=0.0071
	AGG	0.30 (225)	0.31 (191)	p=0.7294

Amino Acids	Codons	Domain	Non-Domain	Chi-Sqaured
Ser	UCU	0.11 (185)	0.10 (130)	p=0.2342
	UCC	0.23 (372)	0.19 (244)	p=0.0070
	UCA	0.12 (200)	0.12 (152)	p=0.6190
	UCG	0.14 (223)	0.11 (142)	p=0.0236
	AGU	0.15 (243)	0.19 (255)	p=0.0008
	AGC	0.26 (425)	0.30 (394)	p=0.0125
Thr	ACU	0.15 (185)	0.19 (141)	p=0.0184
	ACC	0.32 (397)	0.30 (225)	p=0.4214
	ACA	0.24 (305)	0.23 (175)	p=0.5998
	ACG	0.29 (363)	0.28 (208)	p=0.5430
Tyr	UAU	0.33 (295)	0.38 (198)	p=0.0471
	UAC	0.67 (599)	0.62 (320)	p=0.0471
Ala	GCU	0.16 (316)	0.16 (183)	p=0.9863
	GCC	0.24 (456)	0.30 (339)	p<0.0001
	GCA*	0.42 (809)	0.29 (329)	p<0.0001
	GCG	0.18 (345)	0.24 (266)	p<0.0001
Ile	AUU	0.36 (559)	0.36 (233)	p=0.8516
	AUC	0.32 (498)	0.33 (214)	p=0.7791
	AUA	0.31 (482)	0.31 (202)	p=0.9287
Asp	GAU	0.47 (692)	0.49 (552)	p=0.2788
	GAC	0.53 (776)	0.51 (568)	p=0.2788
Glu	GAA*	0.64 (1375)	0.57 (899)	p<0.0001
	GAG	0.36 (783)	0.43 (677)	p<0.0001
Gly	GGU	0.16 (225)	0.18 (250)	p=0.1834
	GGC	0.23 (313)	0.25 (348)	p=0.0996
	GGA	0.38 (526)	0.34 (466)	p=0.0234
	GGG	0.23 (319)	0.23 (311)	p=0.7795

Table 96 | Codon Usage from Domains and Non-Domains for the Alveolate *P. vivax*

Amino Acids	Codons	Domain	Non-Domain	Chi-Sqaured	Amino Acids	Codons	Domain	Non-Domain	Chi-Sqaured
Phe	UUU	0.16 (90)	0.22 (41)	p=0.0678	**Ser**	UCU	0.11 (103)	0.13 (49)	p=0.3874
	UUC	0.84 (458)	0.78 (142)	p=0.0678		UCC	0.15 (144)	0.21 (81)	p=0.0144
Leu	UUA	0.01 (6)	0.01 (2)	p=0.8000		UCA	0.04 (36)	0.04 (14)	p=0.8448
	UUG	0.08 (95)	0.10 (30)	p=0.4479		UCG	0.48 (449)	0.38 (150)	p=0.0024
	CUU	0.14 (158)	0.10 (32)	p=0.1047		AGU	0.05 (45)	0.05 (20)	p=0.7804
	CUC	0.25 (284)	0.27 (85)	p=0.3582		AGC	0.18 (167)	0.19 (76)	p=0.4394
	CUA	0.03 (31)	0.05 (17)	p=0.0155	**Thr**	ACU	0.17 (180)	0.17 (71)	p=0.8120
	CUG	0.50 (576)	0.47 (146)	p=0.3023		ACC	0.34 (351)	0.31 (129)	p=0.2304
Val	GUU	0.11 (122)	0.17 (67)	p=0.0020		ACA	0.07 (71)	0.08 (35)	p=0.3319
	GUC	0.26 (281)	0.30 (116)	p=0.1098		ACG	0.42 (431)	0.44 (185)	p=0.4163
	GUA	0.03 (35)	0.03 (12)	p=0.9155	**Tyr**	UAU	0.07 (34)	0.07 (10)	p=0.8885
	GUG*	0.60 (652)	0.50 (192)	p=0.0005		UAC	0.93 (455)	0.93 (127)	p=0.8885
Pro	CCU	0.25 (202)	0.24 (77)	p=0.8282	**Ala**	GCU	0.35 (438)	0.36 (249)	p=0.5836
	CCC	0.31 (249)	0.33 (104)	p=0.4947		GCC*	0.51 (641)	0.42 (295)	p=0.0005
	CCA	0.10 (80)	0.07 (22)	p=0.1237		GCA	0.03 (44)	0.06 (42)	p=0.0079
	CCG	0.35 (285)	0.36 (116)	p=0.6489		GCG	0.11 (145)	0.16 (110)	p=0.0059
His	CAU	0.05 (11)	0.12 (12)	p=0.0090	**Ile**	AUU	0.28 (192)	0.20 (35)	p=0.0248
	CAC	0.95 (233)	0.88 (85)	p=0.0090		AUC	0.71 (492)	0.78 (139)	p=0.0896
Gln	CAA	0.08 (46)	0.10 (22)	p=0.3773		AUA	0.01 (6)	0.03 (5)	p=0.0402
	CAG	0.92 (500)	0.90 (188)	p=0.3773	**Asp**	GAU	0.16 (126)	0.17 (49)	p=0.5963
Asn	AAU	0.07 (44)	0.12 (21)	p=0.0532		GAC	0.84 (672)	0.83 (237)	p=0.5963
	AAC	0.93 (552)	0.88 (154)	p=0.0532	**Glu**	GAA	0.13 (129)	0.17 (68)	p=0.1064
Lys	AAA	0.05 (50)	0.08 (33)	p=0.0219		GAG	0.87 (834)	0.83 (338)	p=0.1064
	AAG	0.95 (1041)	0.92 (406)	p=0.0219	**Gly**	GGU*	0.32 (345)	0.22 (92)	p=0.0005
Cys	UGU	0.18 (42)	0.21 (15)	p=0.5332		GGC	0.57 (616)	0.63 (256)	p=0.0384
	UGC	0.82 (190)	0.79 (55)	p=0.5332		GGA	0.09 (103)	0.10 (42)	p=0.6439
Arg	CGU*	0.49 (361)	0.33 (81)	p<0.0001		GGG	0.02 (23)	0.05 (19)	p=0.0083
	CGC	0.47 (343)	0.52 (127)	p=0.1346					
	CGA	0.02 (13)	0.05 (13)	p=0.0027					
	CGG	0.01 (9)	0.04 (9)	p=0.0128					
	AGA	0.01 (4)	0.03 (7)	p=0.0028					
	AGG	0.01 (4)	0.02 (6)	p=0.0098					

Table 97 | Codon Usage from Domains and Non-Domains for the Stramenopile *P. infestans*

Amino Acids	Codons	Domain	Non-Domain	Chi-Sqaured	Amino Acids	Codons	Domain	Non-Domain	Chi-Sqaured
Phe	UUU	0.03 (15)	0.07 (11)	p=0.0348	**Ser**	UCU	0.03 (19)	0.05 (19)	p=0.0351
	UUC	0.97 (461)	0.93 (146)	p=0.0348		UCC	0.19 (141)	0.20 (75)	p=0.8237
Leu	UUA	0.00 (0)	0.00 (1)	p=0.0653		UCA	0.01 (11)	0.02 (8)	p=0.4565
	UUG	0.02 (24)	0.03 (11)	p=0.2149		UCG	0.53 (393)	0.46 (174)	p=0.0154
	CUU	0.05 (51)	0.06 (19)	p=0.3685		AGU	0.02 (14)	0.03 (10)	p=0.4278
	CUC	0.33 (364)	0.38 (122)	p=0.1264		AGC	0.22 (161)	0.25 (96)	p=0.2067
	CUA	0.00 (4)	0.00 (1)	p=0.8832	**Thr**	ACU	0.05 (34)	0.11 (36)	p<0.0001
	CUG	0.59 (647)	0.52 (167)	p=0.0194		ACC	0.43 (314)	0.36 (116)	p=0.0513
Val	GUU	0.04 (39)	0.07 (28)	p=0.0180		ACA	0.01 (8)	0.01 (3)	p=0.8260
	GUC	0.31 (304)	0.34 (136)	p=0.2839		ACG	0.52 (380)	0.52 (165)	p=0.9838
	GUA	0.01 (8)	0.02 (8)	p=0.0625	**Tyr**	UAU	0.01 (4)	0.01 (1)	p=0.9892
	GUG	0.64 (636)	0.57 (231)	p=0.0129		UAC	0.99 (398)	0.99 (98)	p=0.9892
Pro	CCU	0.08 (40)	0.07 (21)	p=0.7741	**Ala**	GCU	0.15 (177)	0.18 (131)	p=0.1271
	CCC	0.52 (264)	0.45 (127)	p=0.0348		GCC*	0.71 (837)	0.63 (470)	p=0.0004
	CCA	0.01 (7)	0.02 (6)	p=0.4477		GCA	0.01 (8)	0.01 (8)	p=0.3493
	CCG	0.38 (193)	0.46 (131)	p=0.0354		GCG	0.13 (154)	0.18 (132)	p=0.0047
His	CAU	0.01 (3)	0.03 (3)	p=0.2675	**Ile**	AUU	0.13 (81)	0.12 (20)	p=0.8743
	CAC	0.99 (226)	0.97 (93)	p=0.2675		AUC	0.87 (561)	0.88 (145)	p=0.7920
Gln	CAA	0.02 (10)	0.04 (9)	p=0.1270		AUA	0.00 (2)	0.00 (0)	p=0.4735
	CAG	0.98 (432)	0.96 (193)	p=0.1270	**Asp**	GAU	0.07 (53)	0.05 (14)	p=0.0824
Asn	AAU	0.04 (18)	0.01 (1)	p=0.0410		GAC	0.93 (657)	0.95 (295)	p=0.0824
	AAC	0.96 (482)	0.99 (169)	p=0.0410	**Glu**	GAA	0.03 (27)	0.08 (36)	p=0.0002
Lys	AAA	0.01 (11)	0.03 (12)	p=0.0105		GAG*	0.97 (776)	0.92 (396)	p=0.0002
	AAG	0.99 (1022)	0.97 (396)	p=0.0105	**Gly**	GGU	0.20 (176)	0.15 (56)	p=0.0734
Cys	UGU	0.09 (21)	0.08 (4)	p=0.8230		GGC	0.74 (669)	0.76 (277)	p=0.6599
	UGC	0.91 (208)	0.92 (45)	p=0.8230		GGA	0.05 (46)	0.08 (30)	p=0.0370
Arg	CGU	0.27 (188)	0.25 (67)	p=0.4535		GGG	0.01 (7)	0.01 (3)	p=0.9417
	CGC	0.70 (487)	0.65 (177)	p=0.1453					
	CGA	0.01 (8)	0.05 (13)	p=0.0005					
	CGG	0.02 (12)	0.03 (7)	p=0.3911					
	AGA	0.00 (1)	0.02 (6)	p=0.0007					
	AGG	0.00 (3)	0.01 (3)	p=0.2309					

Table 98 | Codon Usage from Domains and Non-Domains for the Stramenopile *P. sojae*

Amino Acids	Codons	Domain	Non-Domain	Chi-Sqaured	Amino Acids	Codons	Domain	Non-Domain	Chi-Sqaured
Phe	UUU	0.59 (322)	0.62 (153)	p=0.3321	**Ser**	UCU	0.20 (195)	0.18 (118)	p=0.3306
	UUC	0.41 (228)	0.38 (93)	p=0.3321		UCC	0.11 (112)	0.12 (81)	p=0.5773
Leu	UUA	0.07 (89)	0.13 (66)	p<0.0001		UCA	0.18 (176)	0.17 (114)	p=0.7556
	UUG	0.24 (289)	0.20 (100)	p=0.0835		UCG	0.27 (261)	0.22 (143)	p=0.0254
	CUU	0.27 (327)	0.22 (110)	p=0.0312		AGU	0.11 (106)	0.12 (81)	p=0.3438
	CUC	0.14 (163)	0.16 (78)	p=0.2450		AGC	0.13 (126)	0.18 (117)	p=0.0057
	CUA	0.10 (125)	0.08 (42)	p=0.2242	**Thr**	ACU	0.29 (294)	0.30 (147)	p=0.5028
	CUG	0.17 (209)	0.20 (100)	p=0.1780		ACC	0.18 (182)	0.21 (102)	p=0.1244
Val	GUU	0.25 (280)	0.26 (133)	p=0.7517		ACA	0.21 (220)	0.22 (107)	p=0.7786
	GUC	0.23 (251)	0.26 (134)	p=0.1186		ACG	0.32 (333)	0.27 (130)	p=0.0269
	GUA	0.15 (168)	0.16 (83)	p=0.5808	**Tyr**	UAU	0.33 (158)	0.46 (73)	p=0.0021
	GUG	0.37 (406)	0.31 (160)	p=0.0355		UAC	0.67 (322)	0.54 (84)	p=0.0021
Pro	CCU	0.34 (298)	0.31 (122)	p=0.2312	**Ala**	GCU	0.42 (530)	0.38 (243)	p=0.0903
	CCC	0.19 (167)	0.15 (60)	p=0.0831		GCC	0.26 (328)	0.24 (154)	p=0.3602
	CCA	0.27 (239)	0.32 (126)	p=0.1121		GCA	0.16 (207)	0.21 (135)	p=0.0118
	CCG	0.20 (171)	0.23 (90)	p=0.2085		GCG	0.15 (191)	0.16 (105)	p=0.4700
His	CAU	0.31 (89)	0.50 (93)	p<0.0001	**Ile**	AUU	0.55 (450)	0.55 (158)	p=0.9359
	CAC*	0.69 (194)	0.50 (93)	p<0.0001		AUC	0.41 (335)	0.35 (101)	p=0.0748
Gln	CAA	0.38 (220)	0.52 (146)	p<0.0001		AUA	0.04 (34)	0.10 (30)	p<0.0001
	CAG*	0.62 (364)	0.48 (133)	p<0.0001	**Asp**	GAU	0.42 (370)	0.45 (187)	p=0.3264
Asn	AAU	0.44 (290)	0.44 (118)	p=0.9141		GAC	0.58 (501)	0.55 (225)	p=0.3264
	AAC	0.56 (372)	0.56 (149)	p=0.9141	**Glu**	GAA	0.42 (416)	0.47 (224)	p=0.1357
Lys	AAA	0.34 (407)	0.37 (206)	p=0.1465		GAG	0.58 (564)	0.53 (257)	p=0.1357
	AAG	0.66 (794)	0.63 (344)	p=0.1465	**Gly**	GGU*	0.43 (474)	0.33 (147)	p=0.0002
Cys	UGU	0.47 (100)	0.40 (34)	p=0.2763		GGC	0.32 (350)	0.33 (145)	p=0.8044
	UGC	0.53 (113)	0.60 (51)	p=0.2763		GGA	0.20 (220)	0.24 (109)	p=0.0549
Arg	CGU*	0.51 (419)	0.30 (112)	p<0.0001		GGG	0.05 (52)	0.10 (44)	p=0.0002
	CGC	0.33 (270)	0.37 (136)	p=0.2121					
	CGA	0.08 (64)	0.13 (47)	p=0.0076					
	CGG	0.03 (25)	0.07 (26)	p=0.0018					
	AGA	0.03 (26)	0.09 (35)	p<0.0001					
	AGG	0.02 (13)	0.04 (14)	p=0.0189					

Table 99 | Codon Usage from Domains and Non-Domains for the Stramenopile *P. halstedii*

Amino Acids	Codons	Domain	Non-Domain	Chi-Sqaured	Amino Acids	Codons	Domain	Non-Domain	Chi-Sqaured
Phe	UUU	0.37 (337)	0.42 (177)	p=0.1265	**Ser**	UCU	0.07 (89)	0.11 (86)	p=0.0078
	UUC	0.63 (565)	0.58 (247)	p=0.1265		UCC	0.19 (233)	0.21 (169)	p=0.2909
Leu	UUA	0.01 (11)	0.02 (21)	p<0.0001		UCA	0.03 (35)	0.07 (54)	p<0.0001
	UUG*	0.42 (709)	0.29 (243)	p<0.0001		UCG*	0.45 (552)	0.35 (281)	p<0.0001
	CUU	0.06 (110)	0.09 (75)	p=0.0289		AGU	0.06 (71)	0.07 (60)	p=0.1422
	CUC	0.31 (533)	0.36 (305)	p=0.0184		AGC	0.20 (238)	0.19 (153)	p=0.7864
	CUA	0.01 (22)	0.03 (24)	p=0.0060	**Thr**	ACU	0.11 (122)	0.12 (88)	p=0.3541
	CUG	0.18 (311)	0.21 (177)	p=0.1157		ACC	0.32 (369)	0.36 (262)	p=0.1061
Val	GUU	0.10 (147)	0.11 (96)	p=0.3761		ACA	0.08 (97)	0.11 (82)	p=0.0484
	GUC	0.43 (618)	0.45 (383)	p=0.2394		ACG*	0.49 (557)	0.41 (299)	p=0.0010
	GUA	0.03 (43)	0.05 (44)	p=0.0071	**Tyr**	UAU	0.10 (67)	0.18 (53)	p=0.0012
	GUG	0.44 (636)	0.38 (322)	p=0.0055		UAC	0.90 (580)	0.82 (242)	p=0.0012
Pro	CCU	0.19 (162)	0.22 (159)	p=0.1145	**Ala**	GCU	0.14 (231)	0.18 (255)	p=0.0038
	CCC	0.37 (313)	0.36 (256)	p=0.7034		GCC	0.42 (684)	0.41 (587)	p=0.7905
	CCA	0.09 (77)	0.13 (94)	p=0.0093		GCA	0.08 (127)	0.10 (139)	p=0.0476
	CCG	0.35 (293)	0.28 (200)	p=0.0064		GCG	0.36 (595)	0.31 (440)	p=0.0017
His	CAU	0.21 (87)	0.27 (76)	p=0.0840	**Ile**	AUU	0.30 (318)	0.29 (122)	p=0.8074
	CAC	0.79 (328)	0.73 (210)	p=0.0840		AUC	0.69 (736)	0.69 (285)	p=0.7479
Gln	CAA	0.58 (394)	0.58 (290)	p=0.9294		AUA	0.01 (7)	0.02 (9)	p=0.0120
	CAG	0.42 (285)	0.42 (212)	p=0.9294	**Asp**	GAU	0.18 (207)	0.23 (169)	p=0.0208
Asn	AAU	0.08 (68)	0.16 (72)	p<0.0001		GAC	0.82 (920)	0.77 (574)	p=0.0208
	AAC*	0.92 (738)	0.84 (383)	p<0.0001	**Glu**	GAA	0.64 (733)	0.65 (457)	p=0.5532
Lys	AAA	0.13 (185)	0.22 (182)	p<0.0001		GAG	0.36 (412)	0.35 (242)	p=0.5532
	AAG*	0.87 (1207)	0.78 (631)	p<0.0001	**Gly**	GGU	0.25 (328)	0.20 (150)	p=0.0212
Cys	UGU	0.12 (40)	0.25 (50)	p<0.0001		GGC	0.50 (654)	0.51 (377)	p=0.5032
	UGC*	0.88 (294)	0.75 (151)	p<0.0001		GGA	0.07 (97)	0.12 (88)	p=0.0005
Arg	CGU	0.19 (176)	0.17 (83)	p=0.2619		GGG	0.18 (241)	0.17 (123)	p=0.3643
	CGC*	0.61 (563)	0.47 (232)	p<0.0001					
	CGA	0.09 (79)	0.17 (84)	p<0.0001					
	CGG	0.09 (83)	0.11 (53)	p=0.3139					
	AGA	0.01 (12)	0.03 (13)	p=0.0732					
	AGG	0.01 (6)	0.06 (31)	p<0.0001					

Table 100 | Codon Usage from Domains and Non-Domains for the Stramenopile *A. invadans*

Amino Acids	Codons	Domain	Non-Domain	Chi-Sqaured	Amino Acids	Codons	Domain	Non-Domain	Chi-Sqaured
Phe	UUU	0.28 (127)	0.36 (40)	p=0.1047	**Ser**	UCU	0.03 (20)	0.04 (14)	p=0.3053
	UUC	0.72 (328)	0.64 (72)	p=0.1047		UCC	0.08 (52)	0.12 (39)	p=0.0366
Leu	UUA	0.03 (26)	0.05 (16)	p=0.0201		UCA	0.01 (5)	0.09 (29)	p<0.0001
	UUG	0.13 (127)	0.14 (44)	p=0.4005		UCG*	0.70 (449)	0.48 (151)	p<0.0001
	CUU	0.09 (93)	0.13 (39)	p=0.0686		AGU	0.02 (15)	0.07 (22)	p=0.0005
	CUC*	0.70 (706)	0.48 (146)	p<0.0001		AGC	0.15 (99)	0.19 (61)	p=0.1351
	CUA	0.00 (5)	0.03 (9)	p=0.0003	**Thr**	ACU	0.03 (21)	0.11 (29)	p<0.0001
	CUG	0.06 (56)	0.17 (52)	p<0.0001		ACC	0.28 (183)	0.28 (72)	p=0.9909
Val	GUU	0.05 (42)	0.10 (24)	p=0.0028		ACA	0.03 (18)	0.09 (22)	p<0.0001
	GUC	0.73 (644)	0.64 (156)	p=0.0052		ACG*	0.66 (427)	0.52 (132)	p<0.0001
	GUA	0.00 (2)	0.03 (8)	p<0.0001	**Tyr**	UAU	0.06 (22)	0.16 (18)	p=0.0012
	GUG	0.22 (190)	0.23 (55)	p=0.7401		UAC	0.94 (345)	0.84 (98)	p=0.0012
Pro	CCU	0.07 (32)	0.09 (26)	p=0.1549	**Ala**	GCU	0.11 (116)	0.15 (93)	p=0.0052
	CCC	0.38 (187)	0.36 (99)	p=0.4821		GCC*	0.65 (709)	0.52 (318)	p<0.0001
	CCA	0.04 (20)	0.07 (20)	p=0.0621		GCA	0.02 (18)	0.07 (40)	p<0.0001
	CCG	0.51 (251)	0.48 (133)	p=0.3676		GCG	0.23 (245)	0.26 (156)	p=0.1395
His	CAU	0.10 (23)	0.16 (16)	p=0.1578	**Ile**	AUU	0.20 (139)	0.26 (37)	p=0.1014
	CAC	0.90 (202)	0.84 (86)	p=0.1578		AUC	0.79 (560)	0.67 (96)	p=0.0012
Gln	CAA	0.11 (48)	0.32 (49)	p<0.0001		AUA	0.01 (9)	0.08 (11)	p<0.0001
	CAG*	0.89 (371)	0.68 (106)	p<0.0001	**Asp**	GAU	0.15 (103)	0.17 (43)	p=0.5932
Asn	AAU	0.07 (34)	0.17 (24)	p=0.0002		GAC	0.85 (562)	0.83 (211)	p=0.5932
	AAC*	0.93 (473)	0.83 (121)	p=0.0002	**Glu**	GAA	0.28 (190)	0.37 (100)	p=0.0061
Lys	AAA	0.06 (61)	0.12 (48)	p=0.0002		GAG	0.72 (495)	0.63 (172)	p=0.0061
	AAG*	0.94 (912)	0.88 (341)	p=0.0002	**Gly**	GGU	0.28 (224)	0.29 (71)	p=0.7973
Cys	UGU	0.06 (9)	0.09 (4)	p=0.4259		GGC	0.69 (558)	0.63 (157)	p=0.0807
	UGC	0.94 (144)	0.91 (39)	p=0.4259		GGA	0.01 (8)	0.04 (9)	p=0.0039
Arg	CGU	0.13 (85)	0.18 (55)	p=0.0587		GGG	0.02 (16)	0.04 (11)	p=0.0327
	CGC*	0.80 (512)	0.54 (166)	p<0.0001					
	CGA	0.02 (13)	0.10 (31)	p<0.0001					
	CGG	0.03 (21)	0.06 (19)	p=0.0368					
	AGA	0.01 (9)	0.07 (20)	p<0.0001					
	AGG	0.00 (1)	0.05 (16)	p<0.0001					

Table 101 | Codon Usage from Domains and Non-Domains for the Stramenopile *S. parasitica*

Amino Acids	Codons	Domain	Non-Domain	Chi-Sqaured	Amino Acids	Codons	Domain	Non-Domain	Chi-Sqaured
Phe	UUU	0.21 (111)	0.24 (41)	p=0.4732	**Ser**	UCU	0.17 (140)	0.18 (66)	p=0.7753
	UUC	0.79 (418)	0.76 (133)	p=0.4732		UCC	0.39 (315)	0.30 (109)	p=0.0022
Leu	UUA	0.01 (9)	0.01 (5)	p=0.2990		UCA	0.03 (27)	0.05 (17)	p=0.2780
	UUG	0.26 (276)	0.22 (74)	p=0.1582		UCG	0.15 (123)	0.21 (77)	p=0.0145
	CUU	0.32 (342)	0.31 (105)	p=0.7909		AGU	0.11 (89)	0.08 (29)	p=0.1014
	CUC	0.28 (300)	0.33 (111)	p=0.0823		AGC	0.14 (113)	0.19 (68)	p=0.0444
	CUA	0.03 (31)	0.04 (12)	p=0.5349	**Thr**	ACU	0.20 (131)	0.22 (49)	p=0.5551
	CUG	0.10 (110)	0.09 (29)	p=0.3717		ACC	0.56 (365)	0.55 (123)	p=0.8151
Val	GUU	0.34 (295)	0.31 (77)	p=0.4275		ACA	0.10 (64)	0.10 (23)	p=0.8350
	GUC	0.46 (397)	0.47 (117)	p=0.6186		ACG	0.14 (94)	0.13 (29)	p=0.5955
	GUA	0.08 (72)	0.09 (21)	p=0.9058	**Tyr**	UAU	0.23 (68)	0.19 (13)	p=0.4402
	GUG	0.12 (107)	0.13 (32)	p=0.7780		UAC	0.77 (230)	0.81 (57)	p=0.4402
Pro	CCU	0.48 (215)	0.42 (90)	p=0.1452	**Ala**	GCU	0.23 (210)	0.24 (140)	p=0.5895
	CCC	0.31 (138)	0.31 (68)	p=0.8174		GCC	0.58 (527)	0.52 (303)	p=0.0448
	CCA	0.14 (63)	0.15 (33)	p=0.6523		GCA	0.11 (97)	0.14 (82)	p=0.0394
	CCG	0.08 (35)	0.12 (25)	p=0.1072		GCG	0.09 (82)	0.09 (55)	p=0.7287
His	CAU	0.30 (52)	0.56 (28)	p=0.0007	**Ile**	AUU	0.31 (216)	0.32 (44)	p=0.7225
	CAC	0.70 (122)	0.44 (22)	p=0.0007		AUC	0.66 (463)	0.62 (84)	p=0.3367
Gln	CAA	0.55 (197)	0.59 (57)	p=0.4461		AUA	0.03 (22)	0.06 (8)	p=0.1152
	CAG	0.45 (161)	0.41 (39)	p=0.4461	**Asp**	GAU	0.45 (365)	0.47 (98)	p=0.5857
Asn	AAU	0.23 (93)	0.25 (29)	p=0.7352		GAC	0.55 (446)	0.53 (110)	p=0.5857
	AAC	0.77 (310)	0.75 (89)	p=0.7352	**Glu**	GAA	0.58 (449)	0.55 (161)	p=0.2605
Lys	AAA	0.19 (132)	0.18 (53)	p=0.6308		GAG	0.42 (320)	0.45 (134)	p=0.2605
	AAG	0.81 (562)	0.82 (246)	p=0.6308	**Gly**	GGU	0.28 (285)	0.22 (75)	p=0.0444
Cys	UGU	0.21 (31)	0.24 (9)	p=0.7140		GGC	0.17 (171)	0.30 (103)	p<0.0001
	UGC	0.79 (117)	0.76 (29)	p=0.7140		GGA	0.48 (501)	0.41 (139)	p=0.0147
Arg	CGU	0.29 (131)	0.22 (31)	p=0.1190		GGG	0.08 (79)	0.07 (24)	p=0.7206
	CGC	0.19 (86)	0.26 (36)	p=0.0829					
	CGA	0.24 (109)	0.17 (24)	p=0.0884					
	CGG	0.07 (32)	0.09 (12)	p=0.5475					
	AGA	0.12 (53)	0.10 (14)	p=0.5841					
	AGG	0.09 (43)	0.16 (23)	p=0.0220					

Table 102 | Codon Usage from Domains and Non-Domains for the Stramenopile *P. multistriata*

Amino Acids	Codons	Domain	Non-Domain	Chi-Sqaured
Phe	UUU	0.46 (242)	0.51 (194)	p=0.1361
	UUC	0.54 (288)	0.49 (189)	p=0.1361
Leu	UUA	0.07 (97)	0.07 (60)	p=0.7851
	UUG	0.32 (438)	0.30 (271)	p=0.5028
	CUU*	0.24 (329)	0.17 (149)	p<0.0001
	CUC	0.23 (317)	0.26 (232)	p=0.0963
	CUA	0.06 (88)	0.09 (76)	p=0.0532
	CUG	0.08 (109)	0.11 (102)	p=0.0045
Val	GUU	0.32 (358)	0.28 (198)	p=0.0910
	GUC	0.26 (299)	0.24 (168)	p=0.1896
	GUA	0.13 (150)	0.16 (115)	p=0.0785
	GUG	0.29 (329)	0.32 (231)	p=0.1129
Pro	CCU	0.33 (204)	0.27 (150)	p=0.0231
	CCC	0.28 (174)	0.26 (148)	p=0.5399
	CCA	0.28 (174)	0.32 (178)	p=0.1577
	CCG	0.11 (69)	0.15 (84)	p=0.0469
His	CAU	0.45 (143)	0.49 (111)	p=0.3153
	CAC	0.55 (175)	0.51 (114)	p=0.3153
Gln	CAA	0.54 (225)	0.58 (236)	p=0.2968
	CAG	0.46 (192)	0.42 (174)	p=0.2968
Asn	AAU	0.43 (236)	0.42 (217)	p=0.7398
	AAC	0.57 (311)	0.58 (298)	p=0.7398
Lys	AAA	0.27 (250)	0.32 (212)	p=0.0735
	AAG	0.73 (662)	0.68 (460)	p=0.0735
Cys	UGU	0.53 (142)	0.57 (85)	p=0.3810
	UGC	0.47 (128)	0.43 (64)	p=0.3810
Arg	CGU*	0.36 (298)	0.23 (143)	p<0.0001
	CGC	0.15 (123)	0.12 (78)	p=0.1976
	CGA	0.12 (104)	0.18 (110)	p=0.0071
	CGG	0.04 (33)	0.06 (40)	p=0.0365
	AGA	0.16 (132)	0.20 (124)	p=0.0523
	AGG	0.17 (143)	0.21 (133)	p=0.0525
Ser	UCU	0.21 (228)	0.16 (185)	p=0.0050
	UCC	0.21 (226)	0.19 (209)	p=0.1667
	UCA	0.15 (166)	0.16 (176)	p=0.8578
	UCG	0.13 (145)	0.15 (174)	p=0.1729
	AGU	0.20 (212)	0.18 (208)	p=0.4965
	AGC	0.10 (107)	0.16 (177)	p<0.0001
Thr	ACU	0.32 (247)	0.27 (204)	p=0.0597
	ACC	0.29 (229)	0.24 (184)	p=0.0362
	ACA	0.19 (151)	0.27 (201)	p=0.0005
	ACG	0.20 (157)	0.22 (163)	p=0.4260
Tyr	UAU	0.32 (136)	0.41 (115)	p=0.0167
	UAC	0.68 (286)	0.59 (165)	p=0.0167
Ala	GCU	0.36 (436)	0.32 (334)	p=0.0160
	GCC	0.28 (329)	0.26 (277)	p=0.4952
	GCA	0.25 (296)	0.30 (312)	p=0.0105
	GCG	0.11 (135)	0.13 (133)	p=0.3391
Ile	AUU	0.47 (380)	0.40 (171)	p=0.0173
	AUC	0.43 (342)	0.42 (178)	p=0.8095
	AUA	0.10 (79)	0.18 (75)	p<0.0001
Asp	GAU	0.63 (532)	0.58 (398)	p=0.0536
	GAC	0.37 (312)	0.42 (286)	p=0.0536
Glu	GAA	0.40 (369)	0.45 (348)	p=0.0253
	GAG	0.60 (558)	0.55 (422)	p=0.0253
Gly	GGU	0.26 (278)	0.27 (216)	p=0.5679
	GGC	0.13 (141)	0.18 (149)	p=0.0011
	GGA*	0.49 (535)	0.40 (320)	p<0.0001
	GGG	0.12 (136)	0.15 (125)	p=0.0642

Table 103 | Codon Usage from Domains and Non-Domains for the Stramenopile *T. pseudonana*

Amino Acids	Codons	Domain	Non-Domain	Chi-Sqaured
Phe	UUU	0.34 (274)	0.44 (157)	p=0.0021
	UUC	0.66 (528)	0.56 (203)	p=0.0021
Leu	UUA	0.05 (85)	0.06 (49)	p=0.2661
	UUG	0.22 (361)	0.22 (173)	p=0.9144
	CUU	0.15 (237)	0.21 (164)	p<0.0001
	CUC	0.23 (371)	0.24 (182)	p=0.6854
	CUA	0.05 (78)	0.06 (48)	p=0.1460
	CUG*	0.30 (491)	0.20 (155)	p<0.0001
Val	GUU	0.15 (190)	0.16 (86)	p=0.3821
	GUC	0.29 (386)	0.34 (184)	p=0.0354
	GUA	0.09 (122)	0.12 (62)	p=0.1354
	GUG*	0.47 (612)	0.38 (202)	p=0.0005
Pro	CCU	0.27 (195)	0.30 (175)	p=0.1897
	CCC*	0.44 (320)	0.33 (194)	p<0.0001
	CCA	0.08 (60)	0.18 (104)	p<0.0001
	CCG	0.21 (153)	0.19 (109)	p=0.3036
His	CAU	0.32 (120)	0.49 (114)	p<0.0001
	CAC*	0.68 (256)	0.51 (119)	p<0.0001
Gln	CAA	0.35 (205)	0.48 (137)	p=0.0002
	CAG*	0.65 (379)	0.52 (148)	p=0.0002
Asn	AAU	0.34 (221)	0.42 (107)	p=0.0170
	AAC	0.66 (430)	0.58 (145)	p=0.0170
Lys	AAA	0.26 (303)	0.34 (148)	p=0.0026
	AAG	0.74 (854)	0.66 (290)	p=0.0026
Cys	UGU	0.23 (51)	0.46 (72)	p<0.0001
	UGC*	0.77 (175)	0.54 (83)	p<0.0001
Arg	CGU	0.20 (154)	0.13 (79)	p=0.0027
	CGC*	0.37 (292)	0.25 (148)	p<0.0001
	CGA	0.17 (134)	0.15 (90)	p=0.3907
	CGG	0.17 (132)	0.20 (119)	p=0.1004
	AGA	0.04 (30)	0.13 (75)	p<0.0001
	AGG	0.05 (42)	0.13 (75)	p<0.0001
Ser	UCU	0.16 (157)	0.19 (172)	p=0.1992
	UCC*	0.32 (304)	0.22 (199)	p<0.0001
	UCA	0.08 (74)	0.14 (131)	p<0.0001
	UCG	0.19 (185)	0.17 (157)	p=0.1960
	AGU	0.07 (66)	0.11 (98)	p=0.0042
	AGC	0.18 (169)	0.18 (163)	p=0.9905
Thr	ACU	0.18 (181)	0.24 (125)	p=0.0038
	ACC*	0.35 (345)	0.26 (134)	p=0.0010
	ACA	0.14 (135)	0.22 (114)	p<0.0001
	ACG	0.34 (339)	0.27 (139)	p=0.0075
Tyr	UAU	0.28 (146)	0.41 (74)	p=0.0012
	UAC	0.72 (380)	0.59 (108)	p=0.0012
Ala	GCU	0.19 (296)	0.25 (209)	p=0.0006
	GCC*	0.41 (652)	0.33 (281)	p<0.0001
	GCA	0.14 (224)	0.19 (158)	p=0.0040
	GCG	0.26 (403)	0.23 (197)	p=0.2169
Ile	AUU	0.35 (354)	0.38 (103)	p=0.2739
	AUC*	0.59 (607)	0.46 (124)	p<0.0001
	AUA	0.06 (63)	0.16 (43)	p<0.0001
Asp	GAU	0.33 (317)	0.40 (134)	p=0.0115
	GAC	0.67 (652)	0.60 (198)	p=0.0115
Glu	GAA	0.27 (300)	0.40 (153)	p<0.0001
	GAG*	0.73 (793)	0.60 (232)	p<0.0001
Gly	GGU	0.19 (280)	0.20 (128)	p=0.9123
	GGC	0.40 (579)	0.33 (217)	p=0.0027
	GGA	0.18 (264)	0.20 (132)	p=0.2968
	GGG	0.22 (325)	0.27 (178)	p=0.0185

Table 104 | Codon Usage from Domains and Non-Domains for the Stramenopile *N. gaditana*

Amino Acids	Codons	Domain	Non-Domain	Chi-Sqaured	Amino Acids	Codons	Domain	Non-Domain	Chi-Sqaured
Phe	UUU	0.16 (80)	0.17 (71)	p=0.6565	**Ser**	UCU	0.07 (58)	0.05 (50)	p=0.0910
	UUC	0.84 (416)	0.83 (341)	p=0.6565		UCC	0.32 (248)	0.29 (266)	p=0.1933
Leu	UUA	0.02 (27)	0.03 (34)	p=0.2513		UCA	0.07 (53)	0.06 (53)	p=0.3771
	UUG	0.06 (79)	0.08 (95)	p=0.0942		UCG	0.33 (257)	0.34 (313)	p=0.6500
	CUU	0.07 (84)	0.07 (81)	p=0.8728		AGU	0.03 (23)	0.05 (43)	p=0.0671
	CUC	0.62 (775)	0.58 (679)	p=0.0304		AGC	0.18 (142)	0.21 (197)	p=0.1010
	CUA	0.02 (28)	0.04 (46)	p=0.0168	**Thr**	ACU	0.06 (48)	0.09 (60)	p=0.0944
	CUG	0.20 (252)	0.20 (237)	p=0.9907		ACC	0.23 (174)	0.24 (161)	p=0.9220
Val	GUU	0.06 (60)	0.06 (54)	p=0.9180		ACA	0.07 (53)	0.08 (57)	p=0.3829
	GUC	0.61 (598)	0.63 (566)	p=0.3801		ACG	0.63 (469)	0.59 (404)	p=0.1412
	GUA	0.03 (28)	0.03 (30)	p=0.5470	**Tyr**	UAU	0.11 (42)	0.19 (57)	p=0.0052
	GUG	0.30 (290)	0.27 (245)	p=0.2634		UAC	0.89 (336)	0.81 (248)	p=0.0052
Pro	CCU	0.08 (47)	0.07 (64)	p=0.3633	**Ala**	GCU	0.05 (80)	0.05 (100)	p=0.2857
	CCC*	0.44 (248)	0.34 (304)	p<0.0001		GCC	0.39 (586)	0.35 (777)	p=0.0259
	CCA	0.07 (41)	0.05 (45)	p=0.0656		GCA	0.06 (86)	0.05 (109)	p=0.3108
	CCG	0.40 (226)	0.54 (492)	p<0.0001		GCG	0.50 (766)	0.56 (1232)	p=0.0022
His	CAU	0.12 (35)	0.23 (66)	p=0.0002	**Ile**	AUU	0.13 (77)	0.21 (72)	p=0.0009
	CAC*	0.88 (265)	0.77 (218)	p=0.0002		AUC*	0.83 (504)	0.69 (241)	p<0.0001
Gln	CAA	0.11 (52)	0.19 (79)	p=0.0019		AUA	0.05 (29)	0.10 (34)	p=0.0025
	CAG	0.89 (414)	0.81 (347)	p=0.0019	**Asp**	GAU	0.14 (123)	0.13 (152)	p=0.4240
Asn	AAU	0.17 (71)	0.19 (65)	p=0.4199		GAC	0.86 (773)	0.87 (1060)	p=0.4240
	AAC	0.83 (350)	0.81 (275)	p=0.4199	**Glu**	GAA	0.11 (103)	0.15 (192)	p=0.0058
Lys	AAA	0.10 (94)	0.10 (89)	p=0.7202		GAG	0.89 (809)	0.85 (1053)	p=0.0058
	AAG	0.90 (807)	0.90 (808)	p=0.7202	**Gly**	GGU	0.08 (84)	0.06 (78)	p=0.0261
Cys	UGU	0.09 (26)	0.10 (32)	p=0.8486		GGC	0.76 (789)	0.76 (1013)	p=0.6061
	UGC	0.91 (257)	0.90 (300)	p=0.8486		GGA	0.04 (43)	0.07 (88)	p=0.0113
Arg	CGU	0.05 (45)	0.05 (70)	p=0.8174		GGG	0.11 (116)	0.12 (162)	p=0.5281
	CGC*	0.65 (551)	0.54 (748)	p<0.0001					
	CGA	0.04 (35)	0.06 (78)	p=0.1088					
	CGG	0.21 (181)	0.28 (382)	p=0.0007					
	AGA	0.02 (13)	0.03 (35)	p=0.1117					
	AGG	0.03 (23)	0.05 (64)	p=0.0222					

Table 105 | Codon Usage from Domains and Non-Domains for the Stramenopile *A. anophagefferens*

Amino Acids	Codons	Domain	Non-Domain	Chi-Sqaured
Phe	UUU	0.21 (130)	0.50 (74)	p<0.0001
	UUC*	0.79 (497)	0.50 (74)	p<0.0001
Leu	UUA	0.04 (62)	0.23 (83)	p<0.0001
	UUG	0.10 (153)	0.20 (73)	p<0.0001
	CUU	0.41 (606)	0.17 (60)	p<0.0001
	CUC*	0.28 (416)	0.12 (43)	p<0.0001
	CUA	0.07 (101)	0.16 (58)	p<0.0001
	CUG	0.10 (145)	0.12 (43)	p=0.2230
Val	GUU*	0.42 (509)	0.24 (51)	p<0.0001
	GUC	0.27 (329)	0.17 (37)	p=0.0025
	GUA	0.10 (127)	0.31 (66)	p<0.0001
	GUG	0.21 (255)	0.28 (61)	p=0.0148
Pro	CCU*	0.61 (496)	0.26 (37)	p<0.0001
	CCC	0.09 (76)	0.28 (40)	p<0.0001
	CCA	0.22 (182)	0.38 (55)	p<0.0001
	CCG	0.07 (60)	0.09 (13)	p=0.5047
His	CAU	0.26 (67)	0.63 (79)	p<0.0001
	CAC*	0.74 (194)	0.37 (47)	p<0.0001
Gln	CAA	0.36 (230)	0.71 (92)	p<0.0001
	CAG*	0.64 (413)	0.29 (37)	p<0.0001
Asn	AAU	0.29 (190)	0.62 (105)	p<0.0001
	AAC*	0.71 (474)	0.38 (65)	p<0.0001
Lys	AAA	0.25 (297)	0.86 (453)	p<0.0001
	AAG*	0.75 (894)	0.14 (73)	p<0.0001
Cys	UGU	0.28 (72)	0.58 (95)	p<0.0001
	UGC*	0.72 (181)	0.42 (69)	p<0.0001
Arg	CGU*	0.33 (289)	0.19 (45)	p<0.0001
	CGC*	0.26 (225)	0.12 (28)	p<0.0001
	CGA	0.11 (95)	0.14 (33)	p=0.1606
	CGG	0.04 (36)	0.12 (29)	p<0.0001
	AGA	0.14 (120)	0.30 (71)	p<0.0001
	AGG	0.13 (110)	0.12 (27)	p=0.6853
Ser	UCU	0.25 (294)	0.21 (71)	p=0.1560
	UCC*	0.35 (423)	0.13 (43)	p<0.0001
	UCA	0.10 (123)	0.21 (70)	p<0.0001
	UCG	0.12 (144)	0.17 (59)	p=0.0108
	AGU	0.04 (53)	0.17 (58)	p<0.0001
	AGC	0.13 (155)	0.11 (38)	p=0.3799
Thr	ACU*	0.33 (345)	0.20 (44)	p=0.0001
	ACC*	0.42 (435)	0.18 (39)	p<0.0001
	ACA	0.15 (156)	0.40 (88)	p<0.0001
	ACG	0.10 (106)	0.22 (48)	p<0.0001
Tyr	UAU	0.23 (113)	0.64 (96)	p<0.0001
	UAC*	0.77 (377)	0.36 (55)	p<0.0001
Ala	GCU*	0.57 (949)	0.30 (56)	p<0.0001
	GCC	0.13 (211)	0.17 (32)	p=0.0746
	GCA	0.24 (406)	0.37 (68)	p=0.0002
	GCG	0.07 (113)	0.16 (30)	p<0.0001
Ile	AUU	0.38 (352)	0.38 (99)	p=0.8602
	AUC*	0.56 (511)	0.28 (73)	p<0.0001
	AUA	0.06 (54)	0.34 (90)	p<0.0001
Asp	GAU	0.59 (547)	0.62 (56)	p=0.5612
	GAC	0.41 (379)	0.38 (34)	p=0.5612
Glu	GAA	0.45 (467)	0.64 (89)	p<0.0001
	GAG*	0.55 (579)	0.36 (49)	p<0.0001
Gly	GGU	0.27 (395)	0.22 (28)	p=0.2254
	GGC	0.18 (273)	0.25 (32)	p=0.0739
	GGA	0.47 (695)	0.33 (42)	p=0.0018
	GGG	0.08 (122)	0.21 (27)	p<0.0001

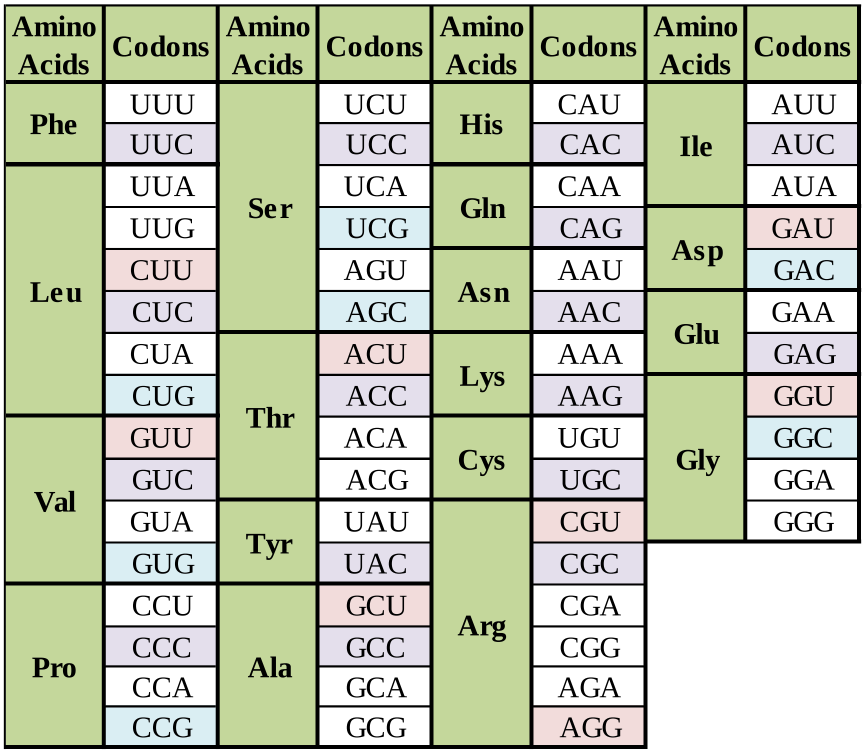
Table 106 | Codon Usage from Domains and Non-Domains for the Rhizarian *B. natans*

Amino Acids	Codons	Domain	Non-Domain	Chi-Sqaured
Phe	UUU	0.08 (48)	0.10 (15)	p=0.4250
	UUC	0.92 (525)	0.90 (128)	p=0.4250
Leu	UUA	0.01 (9)	0.01 (3)	p=0.6528
	UUG	0.07 (95)	0.11 (38)	p=0.0086
	CUU	0.10 (140)	0.10 (35)	p=0.9552
	CUC	0.24 (329)	0.24 (82)	p=0.9481
	CUA	0.02 (24)	0.02 (6)	p=0.9823
	CUG	0.57 (801)	0.53 (182)	p=0.1149
Val	GUU	0.14 (186)	0.14 (52)	p=0.7638
	GUC*	0.63 (859)	0.53 (195)	p=0.0009
	GUA	0.03 (39)	0.05 (20)	p=0.0143
	GUG	0.21 (280)	0.27 (98)	p=0.0094
Pro	CCU	0.13 (86)	0.14 (40)	p=0.8769
	CCC	0.24 (159)	0.23 (68)	p=0.6557
	CCA	0.06 (40)	0.12 (36)	p=0.0015
	CCG	0.56 (367)	0.51 (151)	p=0.1441
His	CAU*	0.20 (69)	0.04 (27)	p<0.0001
	CAC	0.80 (273)	0.96 (664)	p<0.0001
Gln	CAA	0.11 (76)	0.19 (46)	p=0.0017
	CAG	0.89 (613)	0.81 (197)	p=0.0017
Asn	AAU	0.09 (52)	0.16 (23)	p=0.0137
	AAC	0.91 (544)	0.84 (125)	p=0.0137
Lys	AAA	0.10 (127)	0.14 (47)	p=0.0440
	AAG	0.90 (1169)	0.86 (300)	p=0.0440
Cys	UGU	0.16 (35)	0.17 (8)	p=0.8310
	UGC	0.84 (187)	0.83 (39)	p=0.8310
Arg	CGU	0.24 (244)	0.18 (60)	p=0.0248
	CGC	0.47 (473)	0.49 (159)	p=0.6557
	CGA	0.10 (103)	0.11 (35)	p=0.8273
	CGG	0.13 (131)	0.14 (47)	p=0.5492
	AGA	0.01 (15)	0.02 (6)	p=0.6706
	AGG	0.04 (36)	0.06 (20)	p=0.0486
Ser	UCU	0.04 (41)	0.07 (26)	p=0.0190
	UCC	0.25 (238)	0.23 (79)	p=0.4676
	UCA	0.03 (28)	0.05 (18)	p=0.0479
	UCG	0.44 (424)	0.34 (119)	p=0.0016
	AGU	0.03 (31)	0.05 (16)	p=0.2320
	AGC	0.21 (201)	0.26 (89)	p=0.0662
Thr	ACU	0.06 (61)	0.06 (15)	p=0.9389
	ACC	0.32 (326)	0.34 (88)	p=0.4667
	ACA	0.05 (46)	0.10 (26)	p=0.0005
	ACG	0.58 (586)	0.50 (127)	p=0.0229
Tyr	UAU	0.10 (45)	0.18 (17)	p=0.0309
	UAC	0.90 (395)	0.82 (77)	p=0.0309
Ala	GCU	0.11 (173)	0.12 (80)	p=0.4035
	GCC	0.50 (786)	0.47 (308)	p=0.2183
	GCA	0.05 (77)	0.10 (63)	p<0.0001
	GCG	0.35 (549)	0.32 (208)	p=0.1607
Ile	AUU	0.19 (182)	0.18 (26)	p=0.7047
	AUC	0.79 (744)	0.79 (114)	p=0.9761
	AUA	0.02 (19)	0.03 (5)	p=0.2720
Asp	GAU	0.23 (240)	0.24 (72)	p=0.7419
	GAC	0.77 (803)	0.76 (229)	p=0.7419
Glu	GAA	0.26 (270)	0.31 (116)	p=0.0632
	GAG	0.74 (759)	0.69 (255)	p=0.0632
Gly	GGU	0.21 (254)	0.21 (108)	p=0.9980
	GGC	0.59 (717)	0.57 (295)	p=0.4597
	GGA	0.08 (92)	0.10 (51)	p=0.1121
	GGG	0.13 (162)	0.13 (67)	p=0.8364

Table 107 | Codon Usage from Domains and Non-Domains for the Rhizarian *P. brassicae*

6.5 Optimal Codons

General description of tables 108 - 140: Codons shown with a blue background are those determined optimal by CodonW with the use of a correspondence analysis. Codons shown with a red background are those found to be optimal when comparing the top and bottom 5% expressed genes. Codons shown with a purple background are those found to be optimal by both methods.



Amino Acids	Codons	Amino Acids	Codons	Amino Acids	Codons	Amino Acids	Codons
Phe	UUU	**Ser**	UCU	**His**	CAU	**Ile**	AUU
	UUC 🟪		UCC 🟪		CAC 🟪		AUC 🟪
Leu	UUA		UCA	**Gln**	CAA		AUA
	UUG		UCG 🟦		CAG 🟪	**Asp**	GAU 🟥
	CUU 🟥		AGU	**Asn**	AAU		GAC 🟦
	CUC 🟪		AGC 🟦		AAC 🟪	**Glu**	GAA
	CUA	**Thr**	ACU 🟥	**Lys**	AAA		GAG 🟪
	CUG 🟦		ACC 🟪		AAG 🟪	**Gly**	GGU 🟥
Val	GUU 🟥		ACA	**Cys**	UGU		GGC 🟦
	GUC 🟪		ACG		UGC 🟪		GGA
	GUA	**Tyr**	UAU	**Arg**	CGU 🟥		GGG
	GUG 🟦		UAC 🟪		CGC 🟪		
Pro	CCU	**Ala**	GCU 🟥		CGA		
	CCC 🟪		GCC 🟪		CGG		
	CCA		GCA		AGA		
	CCG 🟦		GCG		AGG 🟥		

Table 108 | Optimal Codons for the Amoebozoan *A. castellani*

Amino Acids	Codons	Amino Acids	Codons	Amino Acids	Codons	Amino Acids	Codons
Phe	🟦 UUU	**Ser**	🟦 UCU	**His**	🟦 CAU	**Ile**	🟥 AUU
	🟥 UUC		🟥 UCC		🟥 CAC		🟥 AUC
Leu	🟦 UUA		🟥 UCA	**Gln**	🟥 CAA		🟦 AUA
	UUG		🟦 UCG		🟦 CAG	**Asp**	🟦 GAU
	🟥 CUU		🟦 AGU	**Asn**	🟦 AAU		🟥 GAC
	🟥 CUC		🟥 AGC		🟥 AAC	**Glu**	🟪 GAA
	🟦 CUA	**Thr**	🟥 ACU	**Lys**	🟦 AAA		GAG
	🟦 CUG		🟥 ACC		🟥 AAG	**Gly**	GGU
Val	🟥 GUU		🟦 ACA	**Cys**	🟦 UGU		GGC
	🟥 GUC		🟦 ACG		🟥 UGC		🟥 GGA
	🟦 GUA	**Tyr**	🟦 UAU	**Arg**	🟦 CGU		🟦 GGG
	🟦 GUG		🟥 UAC		CGC		
Pro	🟦 CCU	**Ala**	🟥 GCU		🟦 CGA		
	CCC		🟥 GCC		🟦 CGG		
	🟥 CCA		🟦 GCA		🟥 AGA		
	CCG		🟦 GCG		🟦 AGG		

Table 109 | Optimal Codons for the Amoebozoan *E. histolytica*

Amino Acids	Codons	Amino Acids	Codons	Amino Acids	Codons	Amino Acids	Codons
Phe	🟦 UUU	**Ser**	🟥 UCU	**His**	🟦 CAU	**Ile**	🟦 AUU
	🟥 UUC		🟥 UCC		🟥 CAC		🟥 AUC
Leu	🟦 UUA		🟦 UCA	**Gln**	🟪 CAA		AUA
	UUG		UCG		CAG	**Asp**	🟦 GAU
	🟥 CUU		🟦 AGU	**Asn**	🟦 AAU		🟥 GAC
	🟥 CUC		🟥 AGC		🟥 AAC	**Glu**	🟦 GAA
	CUA	**Thr**	🟪 ACU	**Lys**	🟦 AAA		🟥 GAG
	CUG		🟥 ACC		🟥 AAG	**Gly**	🟪 GGU
Val	🟦 GUU		🟦 ACA	**Cys**	🟦 UGU		🟥 GGC
	🟥 GUC		ACG		🟥 UGC		GGA
	GUA	**Tyr**	🟦 UAU	**Arg**	🟪 CGU		GGG
	GUG		🟥 UAC		CGC		
Pro	CCU	**Ala**	🟪 GCU		CGA		
	CCC		🟥 GCC		CGG		
	🟪 CCA		🟦 GCA		🟦 AGA		
	CCG		GCG		AGG		

Table 110 | Optimal Codons for the Amoebozoan *D. discoideum*

Amino Acids	Codons	Amino Acids	Codons	Amino Acids	Codons	Amino Acids	Codons
Phe	UUU	**Ser**	UCU	**His**	CAU	**Ile**	AUU
	UUC		UCC		CAC		AUC
Leu	UUA		UCA	**Gln**	CAA		AUA
	UUG		UCG		CAG	**Asp**	GAU
	CUU		AGU	**Asn**	AAU		GAC
	CUC		AGC		AAC	**Glu**	GAA
	CUA	**Thr**	ACU	**Lys**	AAA		GAG
	CUG		ACC		AAG	**Gly**	GGU
Val	GUU		ACA	**Cys**	UGU		GGC
	GUC		ACG		UGC		GGA
	GUA	**Tyr**	UAU	**Arg**	CGU		GGG
	GUG		UAC		CGC		
Pro	CCU	**Ala**	GCU		CGA		
	CCC		GCC		CGG		
	CCA		GCA		AGA		
	CCG		GCG		AGG		

Table 111 | Optimal Codons for the Amoebozoan *P. fungivorum*

Amino Acids	Codons	Amino Acids	Codons	Amino Acids	Codons	Amino Acids	Codons
Phe	UUU	**Ser**	UCU	**His**	CAU	**Ile**	AUU
	UUC		UCC		CAC		AUC
Leu	UUA		UCA	**Gln**	CAA		AUA
	UUG		UCG		CAG	**Asp**	GAU
	CUU		AGU	**Asn**	AAU		GAC
	CUC		AGC		AAC	**Glu**	GAA
	CUA	**Thr**	ACU	**Lys**	AAA		GAG
	CUG		ACC		AAG	**Gly**	GGU
Val	GUU		ACA	**Cys**	UGU		GGC
	GUC		ACG		UGC		GGA
	GUA	**Tyr**	UAU	**Arg**	CGU		GGG
	GUG		UAC		CGC		
Pro	CCU	**Ala**	GCU		CGA		
	CCC		GCC		CGG		
	CCA		GCA		AGA		
	CCG		GCG		AGG		

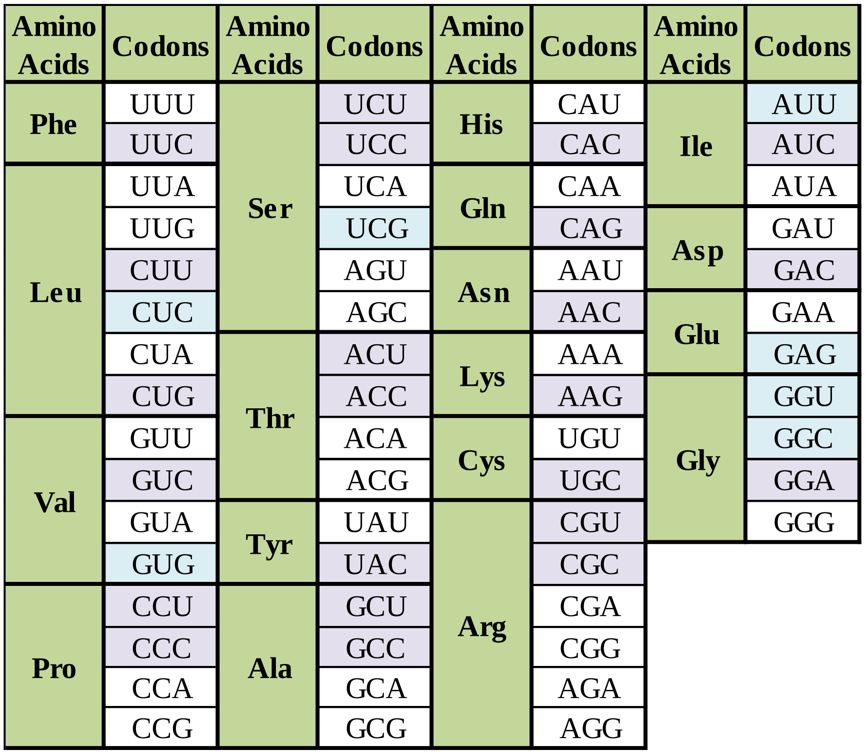
Table 112 | Optimal Codons for the Apusomonadan *T. trahens*

Amino Acids	Codons	Amino Acids	Codons	Amino Acids	Codons	Amino Acids	Codons
Phe	UUU	Ser	UCU	His	CAU	Ile	AUU
	UUC		UCC		CAC		AUC
Leu	UUA		UCA	Gln	CAA		AUA
	UUG		UCG		CAG	Asp	GAU
	CUU		AGU	Asn	AAU		GAC
	CUC		AGC		AAC	Glu	GAA
	CUA	Thr	ACU	Lys	AAA		GAG
	CUG		ACC		AAG	Gly	GGU
Val	GUU		ACA	Cys	UGU		GGC
	GUC		ACG		UGC		GGA
	GUA	Tyr	UAU	Arg	CGU		GGG
	GUG		UAC		CGC		
Pro	CCU	Ala	GCU		CGA		
	CCC		GCC		CGG		
	CCA		GCA		AGA		
	CCG		GCG		AGG		

Table 113 | Optimal Codons for the Opisthokont *F. alba*

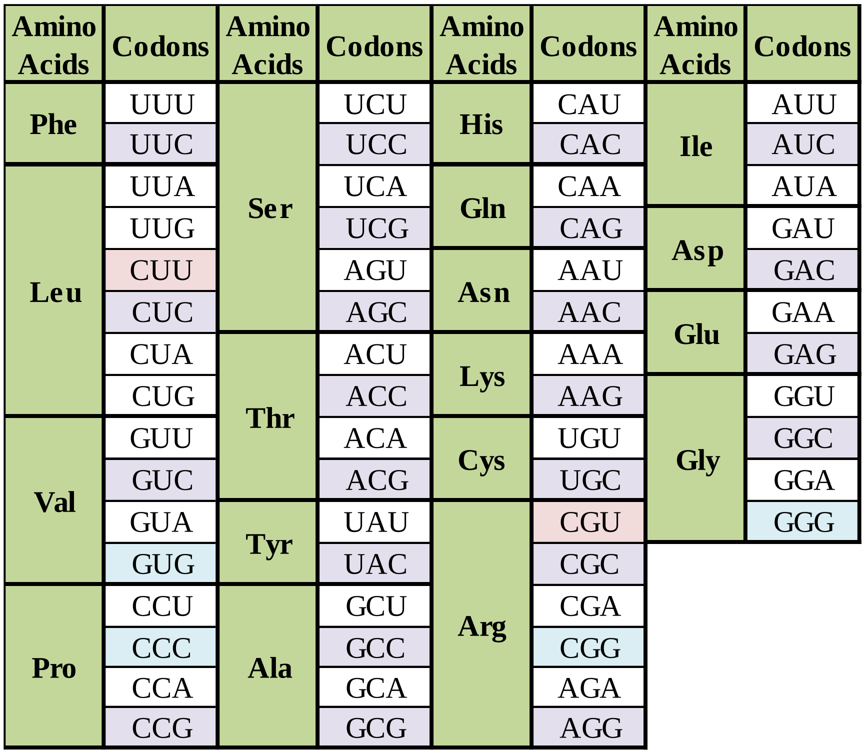
Amino Acids	Codons	Amino Acids	Codons	Amino Acids	Codons	Amino Acids	Codons
Phe	UUU	Ser	UCU	His	CAU	Ile	AUU
	UUC		UCC		CAC		AUC
Leu	UUA		UCA	Gln	CAA		AUA
	UUG		UCG		CAG	Asp	GAU
	CUU		AGU	Asn	AAU		GAC
	CUC		AGC		AAC	Glu	GAA
	CUA	Thr	ACU	Lys	AAA		GAG
	CUG		ACC		AAG	Gly	GGU
Val	GUU		ACA	Cys	UGU		GGC
	GUC		ACG		UGC		GGA
	GUA	Tyr	UAU	Arg	CGU		GGG
	GUG		UAC		CGC		
Pro	CCU	Ala	GCU		CGA		
	CCC		GCC		CGG		
	CCA		GCA		AGA		
	CCG		GCG		AGG		

Table 114 | Optimal Codons for the Opisthokont *C. owczarzaki*



Amino Acids	Codons	Amino Acids	Codons	Amino Acids	Codons	Amino Acids	Codons
Phe	UUU	Ser	UCU	His	CAU	Ile	AUU
	UUC		UCC		CAC		AUC
Leu	UUA		UCA	Gln	CAA		AUA
	UUG		UCG		CAG	Asp	GAU
	CUU		AGU	Asn	AAU		GAC
	CUC		AGC		AAC	Glu	GAA
	CUA	Thr	ACU	Lys	AAA		GAG
	CUG		ACC		AAG	Gly	GGU
Val	GUU		ACA	Cys	UGU		GGC
	GUC		ACG		UGC		GGA
	GUA	Tyr	UAU	Arg	CGU		GGG
	GUG		UAC		CGC		
Pro	CCU	Ala	GCU		CGA		
	CCC		GCC		CGG		
	CCA		GCA		AGA		
	CCG		GCG		AGG		

Table 115 | Optimal Codons for the Opisthokont *S. arctica*



Amino Acids	Codons	Amino Acids	Codons	Amino Acids	Codons	Amino Acids	Codons
Phe	UUU	Ser	UCU	His	CAU	Ile	AUU
	UUC		UCC		CAC		AUC
Leu	UUA		UCA	Gln	CAA		AUA
	UUG		UCG		CAG	Asp	GAU
	CUU		AGU	Asn	AAU		GAC
	CUC		AGC		AAC	Glu	GAA
	CUA	Thr	ACU	Lys	AAA		GAG
	CUG		ACC		AAG	Gly	GGU
Val	GUU		ACA	Cys	UGU		GGC
	GUC		ACG		UGC		GGA
	GUA	Tyr	UAU	Arg	CGU		GGG
	GUG		UAC		CGC		
Pro	CCU	Ala	GCU		CGA		
	CCC		GCC		CGG		
	CCA		GCA		AGA		
	CCG		GCG		AGG		

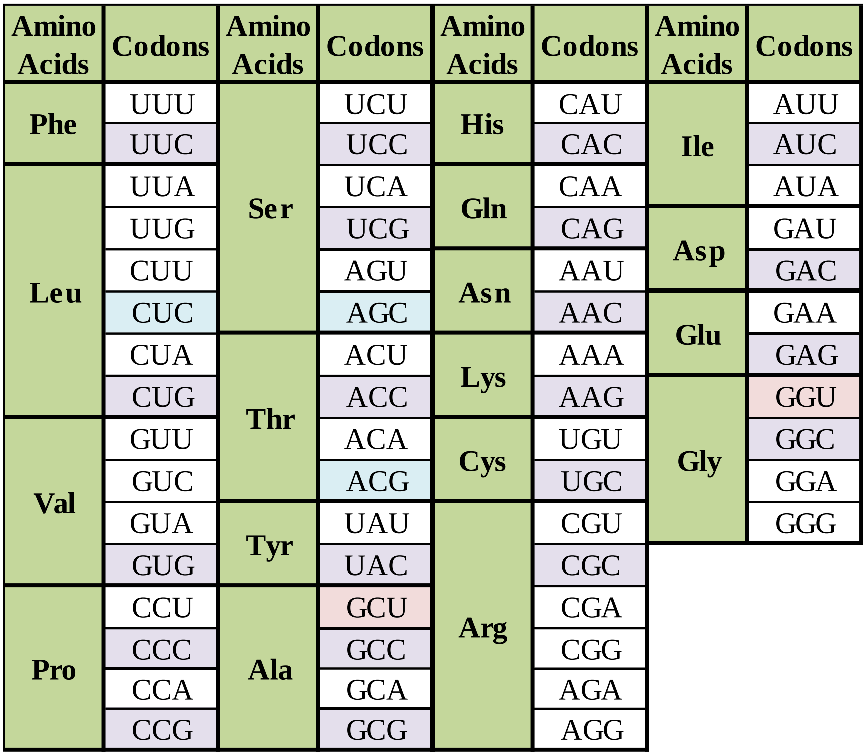
Table 116 | Optimal Codons for the Excavate *G. lamblia*

Amino Acids	Codons	Amino Acids	Codons	Amino Acids	Codons	Amino Acids	Codons
Phe	UUU	**Ser**	UCU	**His**	CAU	**Ile**	AUU
	UUC		UCC		CAC		AUC
Leu	UUA		UCA	**Glu**	GAA		AUA
	UUG		UCG		GAG	**Asp**	GAU
	CUU		AGU	**Asn**	AAU		GAC
	CUC		AGC		AAC	**Gln**	CAA
	CUA	**Thr**	ACU	**Lys**	AAA		CAG
	CUG		ACC		AAG		UAA
Val	GUU		ACA	**Cys**	UGU		UAG
	GUC		ACG		UGC	**Gly**	GGU
	GUA	**Tyr**	UAU	**Arg**	CGU		GGC
	GUG		UAC		CGC		GGA
Pro	CCU	**Ala**	GCU		CGA		GGG
	CCC		GCC		CGG		
	CCA		GCA		AGA		
	CCG		GCG		AGG		

Table 117 | Optimal Codons for the Excavate *S. salmonicida*

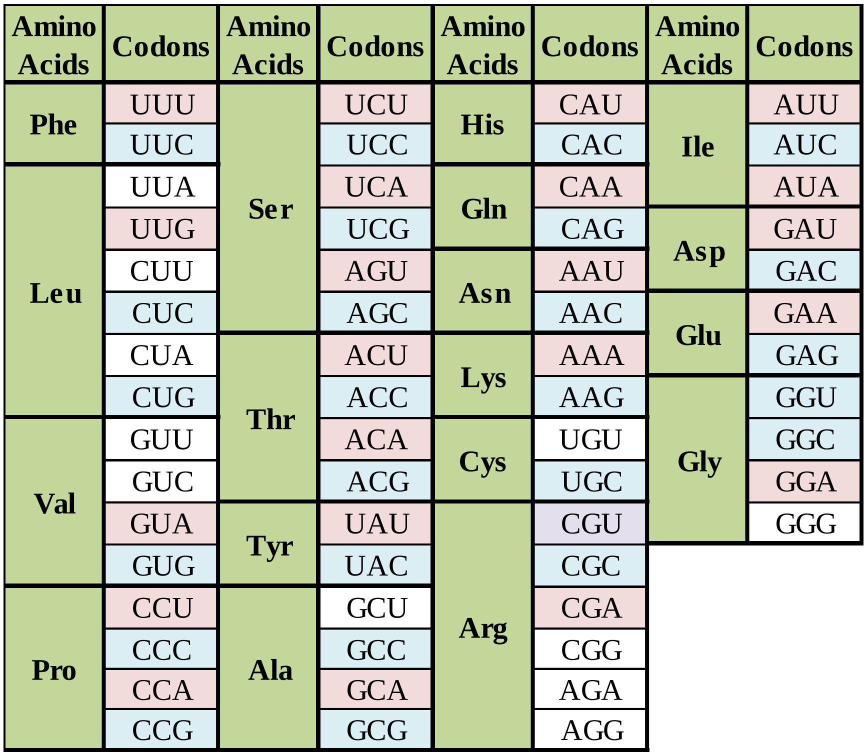
Amino Acids	Codons	Amino Acids	Codons	Amino Acids	Codons	Amino Acids	Codons
Phe	UUU	**Ser**	UCU	**His**	CAU	**Ile**	AUU
	UUC		UCC		CAC		AUC
Leu	UUA		UCA	**Gln**	CAA		AUA
	UUG		UCG		CAG	**Asp**	GAU
	CUU		AGU	**Asn**	AAU		GAC
	CUC		AGC		AAC	**Glu**	GAA
	CUA	**Thr**	ACU	**Lys**	AAA		GAG
	CUG		ACC		AAG	**Gly**	GGU
Val	GUU		ACA	**Cys**	UGU		GGC
	GUC		ACG		UGC		GGA
	GUA	**Tyr**	UAU	**Arg**	CGU		GGG
	GUG		UAC		CGC		
Pro	CCU	**Ala**	GCU		CGA		
	CCC		GCC		CGG		
	CCA		GCA		AGA		
	CCG		GCG		AGG		

Table 118 | Optimal Codons for the Excavate *L. major*



Amino Acids	Codons	Amino Acids	Codons	Amino Acids	Codons	Amino Acids	Codons
Phe	UUU	**Ser**	UCU	**His**	CAU	**Ile**	AUU
	UUC		UCC		CAC		AUC
Leu	UUA		UCA	**Gln**	CAA		AUA
	UUG		UCG		CAG	**Asp**	GAU
	CUU		AGU	**Asn**	AAU		GAC
	CUC		AGC		AAC	**Glu**	GAA
	CUA	**Thr**	ACU	**Lys**	AAA		GAG
	CUG		ACC		AAG	**Gly**	GGU
Val	GUU		ACA	**Cys**	UGU		GGC
	GUC		ACG		UGC		GGA
	GUA	**Tyr**	UAU	**Arg**	CGU		GGG
	GUG		UAC		CGC		
Pro	CCU	**Ala**	GCU		CGA		
	CCC		GCC		CGG		
	CCA		GCA		AGA		
	CCG		GCG		AGG		

Table 119 | Optimal Codons for the Excavate *L. pyrrhocoris*



Amino Acids	Codons	Amino Acids	Codons	Amino Acids	Codons	Amino Acids	Codons
Phe	UUU	**Ser**	UCU	**His**	CAU	**Ile**	AUU
	UUC		UCC		CAC		AUC
Leu	UUA		UCA	**Gln**	CAA		AUA
	UUG		UCG		CAG	**Asp**	GAU
	CUU		AGU	**Asn**	AAU		GAC
	CUC		AGC		AAC	**Glu**	GAA
	CUA	**Thr**	ACU	**Lys**	AAA		GAG
	CUG		ACC		AAG	**Gly**	GGU
Val	GUU		ACA	**Cys**	UGU		GGC
	GUC		ACG		UGC		GGA
	GUA	**Tyr**	UAU	**Arg**	CGU		GGG
	GUG		UAC		CGC		
Pro	CCU	**Ala**	GCU		CGA		
	CCC		GCC		CGG		
	CCA		GCA		AGA		
	CCG		GCG		AGG		

Table 120 | Optimal Codons for the Excavate *L. seymouri*

Amino Acids	Codons	Amino Acids	Codons	Amino Acids	Codons	Amino Acids	Codons
Phe	UUU	**Ser**	UCU	**His**	CAU	**Ile**	AUU
	UUC		UCC		CAC		AUC
Leu	UUA		UCA	**Gln**	CAA		AUA
	UUG		UCG		CAG	**Asp**	GAU
	CUU		AGU	**Asn**	AAU		GAC
	CUC		AGC		AAC	**Glu**	GAA
	CUA	**Thr**	ACU	**Lys**	AAA		GAG
	CUG		ACC		AAG	**Gly**	GGU
Val	GUU		ACA	**Cys**	UGU		GGC
	GUC		ACG		UGC		GGA
	GUA	**Tyr**	UAU	**Arg**	CGU		GGG
	GUG		UAC		CGC		
Pro	CCU	**Ala**	GCU		CGA		
	CCC		GCC		CGG		
	CCA		GCA		AGA		
	CCG		GCG		AGG		

Table 121 | Optimal Codons for the Excavate *T. brucei*

Amino Acids	Codons	Amino Acids	Codons	Amino Acids	Codons	Amino Acids	Codons
Phe	UUU	**Ser**	UCU	**His**	CAU	**Ile**	AUU
	UUC		UCC		CAC		AUC
Leu	UUA		UCA	**Gln**	CAA		AUA
	UUG		UCG		CAG	**Asp**	GAU
	CUU		AGU	**Asn**	AAU		GAC
	CUC		AGC		AAC	**Glu**	GAA
	CUA	**Thr**	ACU	**Lys**	AAA		GAG
	CUG		ACC		AAG	**Gly**	GGU
Val	GUU		ACA	**Cys**	UGU		GGC
	GUC		ACG		UGC		GGA
	GUA	**Tyr**	UAU	**Arg**	CGU		GGG
	GUG		UAC		CGC		
Pro	CCU	**Ala**	GCU		CGA		
	CCC		GCC		CGG		
	CCA		GCA		AGA		
	CCG		GCG		AGG		

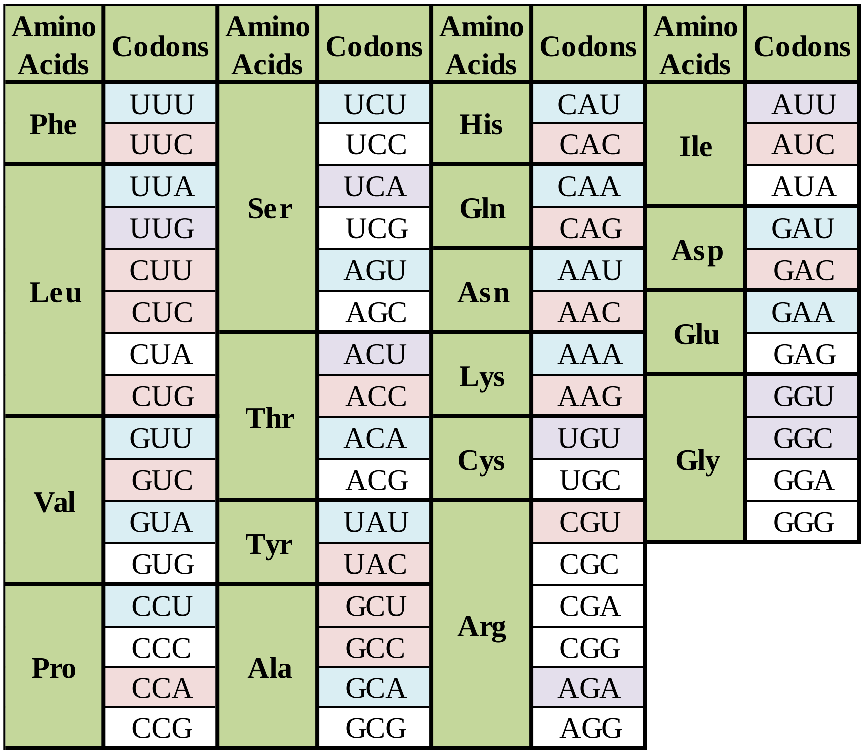
Table 122 | Optimal Codons for the Excavate *T. cruzi*

Amino Acids	**Codons**	**Amino Acids**	**Codons**	**Amino Acids**	**Codons**	**Amino Acids**	**Codons**
Phe	UUU	**Ser**	UCU	**His**	CAU	**Ile**	AUU
	UUC		UCC		CAC		AUC
Leu	UUA		UCA	**Gln**	CAA		AUA
	UUG		UCG		CAG	**Asp**	GAU
	CUU		AGU	**Asn**	AAU		GAC
	CUC		AGC		AAC	**Glu**	GAA
	CUA	**Thr**	ACU	**Lys**	AAA		GAG
	CUG		ACC		AAG	**Gly**	GGU
Val	GUU		ACA	**Cys**	UGU		GGC
	GUC		ACG		UGC		GGA
	GUA	**Tyr**	UAU	**Arg**	CGU		GGG
	GUG		UAC		CGC		
Pro	CCU	**Ala**	GCU		CGA		
	CCC		GCC		CGG		
	CCA		GCA		AGA		
	CCG		GCG		AGG		

Table 123 | Optimal Codons for the Excavate *T. theileri*

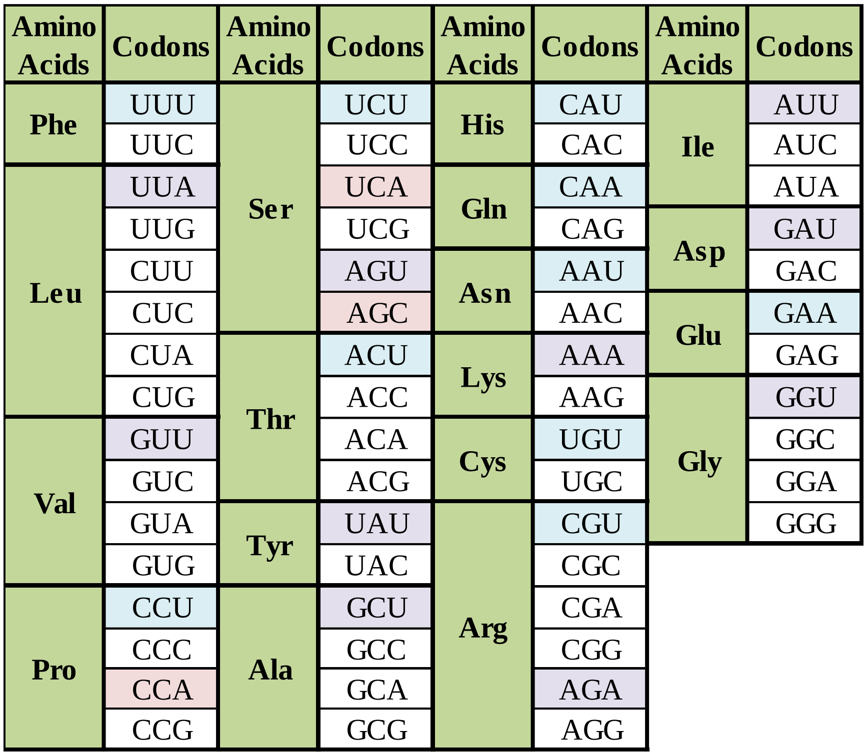
Amino Acids	**Codons**	**Amino Acids**	**Codons**	**Amino Acids**	**Codons**	**Amino Acids**	**Codons**
Phe	UUU	**Ser**	UCU	**His**	CAU	**Ile**	AUU
	UUC		UCC		CAC		AUC
Leu	UUA		UCA	**Glu**	GAA		AUA
	UUG		UCG		GAG	**Asp**	GAU
	CUU		AGU	**Asn**	AAU		GAC
	CUC		AGC		AAC	**Gln**	CAA
	CUA	**Thr**	ACU	**Lys**	AAA		CAG
	CUG		ACC		AAG		UAA
Val	GUU		ACA	**Cys**	UGU		UAG
	GUC		ACG		UGC	**Gly**	GGU
	GUA	**Tyr**	UAU	**Arg**	CGU		GGC
	GUG		UAC		CGC		GGA
Pro	CCU	**Ala**	GCU		CGA		GGG
	CCC		GCC		CGG		
	CCA		GCA		AGA		
	CCG		GCG		AGG		

Table 124 | Optimal Codons for the Alveolate *T. thermophila*



Amino Acids	Codons	Amino Acids	Codons	Amino Acids	Codons	Amino Acids	Codons
Phe	UUU	Ser	UCU	His	CAU	Ile	AUU
	UUC		UCC		CAC		AUC
Leu	UUA		UCA	Gln	CAA		AUA
	UUG		UCG		CAG	Asp	GAU
	CUU		AGU	Asn	AAU		GAC
	CUC		AGC		AAC	Glu	GAA
	CUA	Thr	ACU	Lys	AAA		GAG
	CUG		ACC		AAG	Gly	GGU
Val	GUU		ACA	Cys	UGU		GGC
	GUC		ACG		UGC		GGA
	GUA	Tyr	UAU	Arg	CGU		GGG
	GUG		UAC		CGC		
Pro	CCU	Ala	GCU		CGA		
	CCC		GCC		CGG		
	CCA		GCA		AGA		
	CCG		GCG		AGG		

Table 125 | Optimal Codons for the Alveolate *S. coeruleus*



Amino Acids	Codons	Amino Acids	Codons	Amino Acids	Codons	Amino Acids	Codons
Phe	UUU	Ser	UCU	His	CAU	Ile	AUU
	UUC		UCC		CAC		AUC
Leu	UUA		UCA	Gln	CAA		AUA
	UUG		UCG		CAG	Asp	GAU
	CUU		AGU	Asn	AAU		GAC
	CUC		AGC		AAC	Glu	GAA
	CUA	Thr	ACU	Lys	AAA		GAG
	CUG		ACC		AAG	Gly	GGU
Val	GUU		ACA	Cys	UGU		GGC
	GUC		ACG		UGC		GGA
	GUA	Tyr	UAU	Arg	CGU		GGG
	GUG		UAC		CGC		
Pro	CCU	Ala	GCU		CGA		
	CCC		GCC		CGG		
	CCA		GCA		AGA		
	CCG		GCG		AGG		

Table 126 | Optimal Codons for the Alveolate *P. gallinaceum*

Amino Acids	Codons	Amino Acids	Codons	Amino Acids	Codons	Amino Acids	Codons
Phe	🟦 UUU	**Ser**	UCU	**His**	🟦 CAU	**Ile**	AUU
	🟥 UUC		UCC		🟥 CAC		🟥 AUC
Leu	🟥 UUA		UCA	**Gln**	CAA		🟦 AUA
	UUG		UCG		CAG	**Asp**	🟪 GAU
	CUU		🟪 AGU	**Asn**	🟦 AAU		GAC
	CUC		AGC		🟥 AAC	**Glu**	🟪 GAA
	CUA	**Thr**	ACU	**Lys**	AAA		GAG
	CUG		🟥 ACC		AAG	**Gly**	GGU
Val	GUU		🟦 ACA	**Cys**	🟪 UGU		🟦 GGC
	GUC		🟦 ACG		UGC		🟪 GGA
	🟪 GUA	**Tyr**	🟦 UAU	**Arg**	CGU		🟦 GGG
	🟦 GUG		UAC		CGC		
Pro	🟦 CCU	**Ala**	GCU		CGA		
	🟥 CCC		🟥 GCC		CGG		
	🟥 CCA		🟦 GCA		🟪 AGA		
	🟦 CCG		GCG		🟦 AGG		

Table 127 | Optimal Codons for the Alveolate *P. falciparum*

Amino Acids	Codons	Amino Acids	Codons	Amino Acids	Codons	Amino Acids	Codons
Phe	🟦 UUU	**Ser**	🟦 UCU	**His**	🟦 CAU	**Ile**	🟦 AUU
	🟥 UUC		UCC		🟥 CAC		🟥 AUC
Leu	🟦 UUA		🟪 UCA	**Gln**	🟦 CAA		🟦 AUA
	UUG		UCG		CAG	**Asp**	🟦 GAU
	🟦 CUU		🟦 AGU	**Asn**	🟦 AAU		GAC
	🟥 CUC		AGC		🟥 AAC	**Glu**	🟦 GAA
	🟦 CUA	**Thr**	🟦 ACU	**Lys**	🟦 AAA		GAG
	CUG		ACC		🟥 AAG	**Gly**	🟦 GGU
Val	🟦 GUU		🟦 ACA	**Cys**	🟦 UGU		GGC
	🟥 GUC		ACG		UGC		🟪 GGA
	🟦 GUA	**Tyr**	🟦 UAU	**Arg**	🟦 CGU		GGG
	GUG		🟥 UAC		CGC		
Pro	🟦 CCU	**Ala**	🟦 GCU		🟪 CGA		
	CCC		GCC		CGG		
	🟪 CCA		🟪 GCA		🟪 AGA		
	CCG		GCG		AGG		

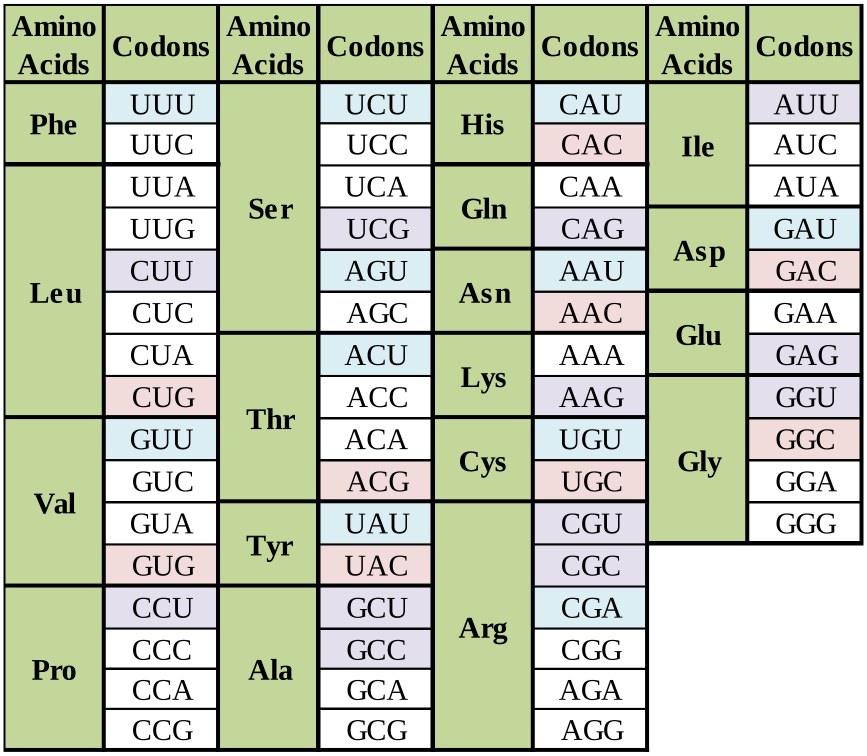
Table 128 | Optimal Codons for the Alveolate *P. vivax*

Amino Acids	Codons	Amino Acids	Codons	Amino Acids	Codons	Amino Acids	Codons
Phe	UUU	**Ser**	UCU	**His**	CAU	**Ile**	AUU
	UUC		UCC		CAC		AUC
Leu	UUA		UCA	**Gln**	CAA		AUA
	UUG		UCG		CAG	**Asp**	GAU
	CUU		AGU	**Asn**	AAU		GAC
	CUC		AGC		AAC	**Glu**	GAA
	CUA	**Thr**	ACU	**Lys**	AAA		GAG
	CUG		ACC		AAG	**Gly**	GGU
Val	GUU		ACA	**Cys**	UGU		GGC
	GUC		ACG		UGC		GGA
	GUA	**Tyr**	UAU	**Arg**	CGU		GGG
	GUG		UAC		CGC		
Pro	CCU	**Ala**	GCU		CGA		
	CCC		GCC		CGG		
	CCA		GCA		AGA		
	CCG		GCG		AGG		

Table 129 | Optimal Codons for the Stramenopile *P. infestans*

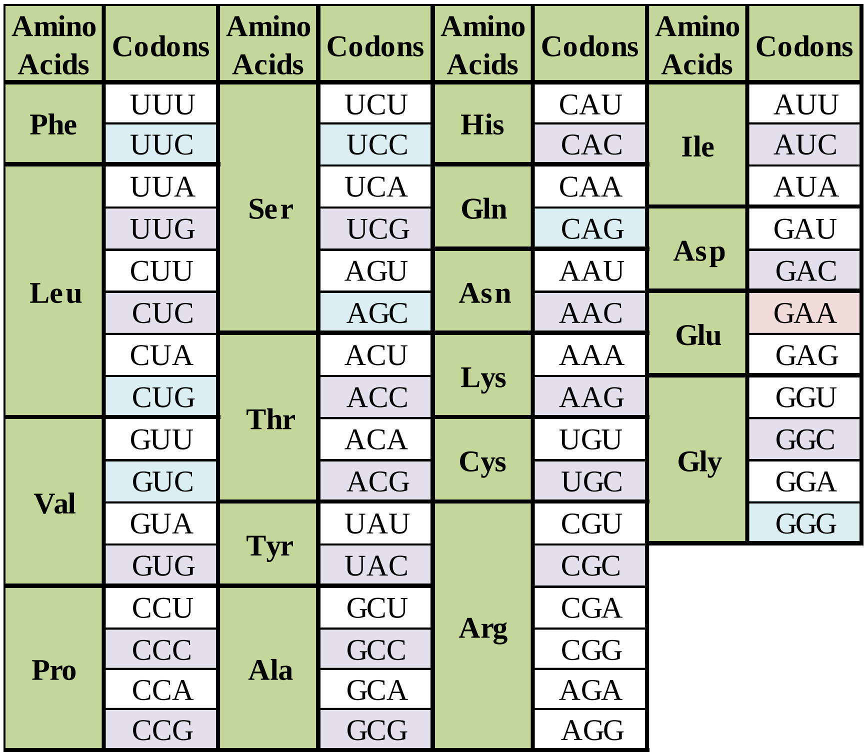
Amino Acids	Codons	Amino Acids	Codons	Amino Acids	Codons	Amino Acids	Codons
Phe	UUU	**Ser**	UCU	**His**	CAU	**Ile**	AUU
	UUC		UCC		CAC		AUC
Leu	UUA		UCA	**Gln**	CAA		AUA
	UUG		UCG		CAG	**Asp**	GAU
	CUU		AGU	**Asn**	AAU		GAC
	CUC		AGC		AAC	**Glu**	GAA
	CUA	**Thr**	ACU	**Lys**	AAA		GAG
	CUG		ACC		AAG	**Gly**	GGU
Val	GUU		ACA	**Cys**	UGU		GGC
	GUC		ACG		UGC		GGA
	GUA	**Tyr**	UAU	**Arg**	CGU		GGG
	GUG		UAC		CGC		
Pro	CCU	**Ala**	GCU		CGA		
	CCC		GCC		CGG		
	CCA		GCA		AGA		
	CCG		GCG		AGG		

Table 130 | Optimal Codons for the Stramenopile *P. sojae*



Amino Acids	Codons	Amino Acids	Codons	Amino Acids	Codons	Amino Acids	Codons
Phe	UUU	Ser	UCU	His	CAU	Ile	AUU
	UUC		UCC		CAC		AUC
Leu	UUA		UCA	Gln	CAA		AUA
	UUG		UCG		CAG	Asp	GAU
	CUU		AGU	Asn	AAU		GAC
	CUC		AGC		AAC	Glu	GAA
	CUA	Thr	ACU	Lys	AAA		GAG
	CUG		ACC		AAG	Gly	GGU
Val	GUU		ACA	Cys	UGU		GGC
	GUC		ACG		UGC		GGA
	GUA	Tyr	UAU	Arg	CGU		GGG
	GUG		UAC		CGC		
Pro	CCU	Ala	GCU		CGA		
	CCC		GCC		CGG		
	CCA		GCA		AGA		
	CCG		GCG		AGG		

Table 131 | Optimal Codons for the Stramenopile *P. halstedii*



Amino Acids	Codons	Amino Acids	Codons	Amino Acids	Codons	Amino Acids	Codons
Phe	UUU	Ser	UCU	His	CAU	Ile	AUU
	UUC		UCC		CAC		AUC
Leu	UUA		UCA	Gln	CAA		AUA
	UUG		UCG		CAG	Asp	GAU
	CUU		AGU	Asn	AAU		GAC
	CUC		AGC		AAC	Glu	GAA
	CUA	Thr	ACU	Lys	AAA		GAG
	CUG		ACC		AAG	Gly	GGU
Val	GUU		ACA	Cys	UGU		GGC
	GUC		ACG		UGC		GGA
	GUA	Tyr	UAU	Arg	CGU		GGG
	GUG		UAC		CGC		
Pro	CCU	Ala	GCU		CGA		
	CCC		GCC		CGG		
	CCA		GCA		AGA		
	CCG		GCG		AGG		

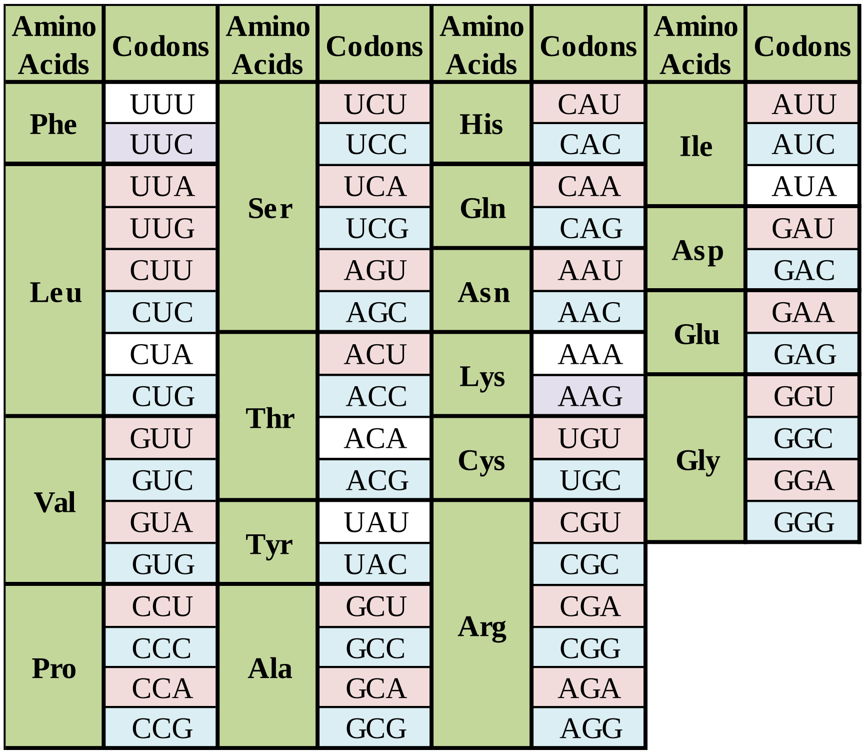
Table 132 | Optimal Codons for the Stramenopile *A. astaci*

Amino Acids	Codons	Amino Acids	Codons	Amino Acids	Codons	Amino Acids	Codons
Phe	UUU	**Ser**	UCU	**His**	CAU	**Ile**	AUU
	UUC		UCC		CAC		AUC
Leu	UUA		UCA	**Gln**	CAA		AUA
	UUG		UCG		CAG	**Asp**	GAU
	CUU		AGU	**Asn**	AAU		GAC
	CUC		AGC		AAC	**Glu**	GAA
	CUA	**Thr**	ACU	**Lys**	AAA		GAG
	CUG		ACC		AAG	**Gly**	GGU
Val	GUU		ACA	**Cys**	UGU		GGC
	GUC		ACG		UGC		GGA
	GUA	**Tyr**	UAU	**Arg**	CGU		GGG
	GUG		UAC		CGC		
Pro	CCU	**Ala**	GCU		CGA		
	CCC		GCC		CGG		
	CCA		GCA		AGA		
	CCG		GCG		AGG		

Table 133 | Optimal Codons for the Stramenopile *A. invadans*

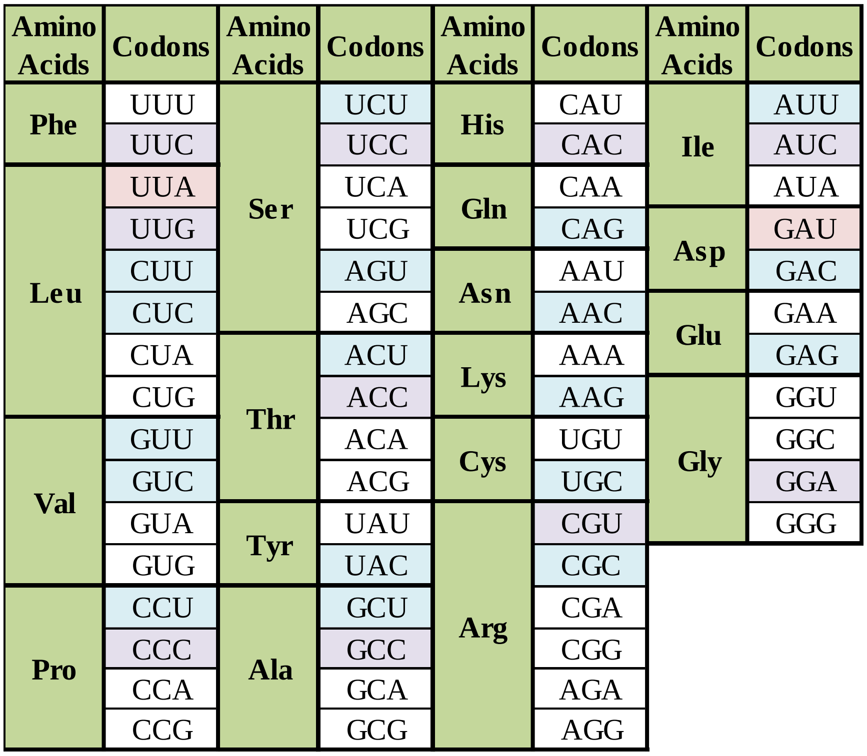
Amino Acids	Codons	Amino Acids	Codons	Amino Acids	Codons	Amino Acids	Codons
Phe	UUU	**Ser**	UCU	**His**	CAU	**Ile**	AUU
	UUC		UCC		CAC		AUC
Leu	UUA		UCA	**Gln**	CAA		AUA
	UUG		UCG		CAG	**Asp**	GAU
	CUU		AGU	**Asn**	AAU		GAC
	CUC		AGC		AAC	**Glu**	GAA
	CUA	**Thr**	ACU	**Lys**	AAA		GAG
	CUG		ACC		AAG	**Gly**	GGU
Val	GUU		ACA	**Cys**	UGU		GGC
	GUC		ACG		UGC		GGA
	GUA	**Tyr**	UAU	**Arg**	CGU		GGG
	GUG		UAC		CGC		
Pro	CCU	**Ala**	GCU		CGA		
	CCC		GCC		CGG		
	CCA		GCA		AGA		
	CCG		GCG		AGG		

Table 134 | Optimal Codons for the Stramenopile *S. parasitica*



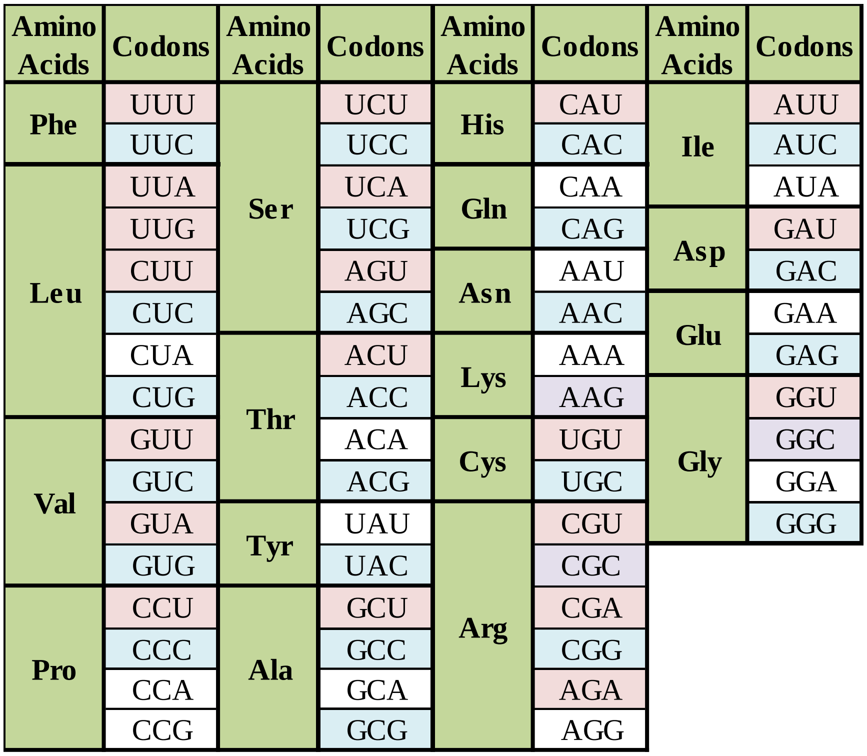
Amino Acids	Codons	Amino Acids	Codons	Amino Acids	Codons	Amino Acids	Codons
Phe	UUU	**Ser**	UCU	**His**	CAU	**Ile**	AUU
	UUC		UCC		CAC		AUC
Leu	UUA		UCA	**Gln**	CAA		AUA
	UUG		UCG		CAG	**Asp**	GAU
	CUU		AGU	**Asn**	AAU		GAC
	CUC		AGC		AAC	**Glu**	GAA
	CUA	**Thr**	ACU	**Lys**	AAA		GAG
	CUG		ACC		AAG	**Gly**	GGU
Val	GUU		ACA	**Cys**	UGU		GGC
	GUC		ACG		UGC		GGA
	GUA	**Tyr**	UAU	**Arg**	CGU		GGG
	GUG		UAC		CGC		
Pro	CCU	**Ala**	GCU		CGA		
	CCC		GCC		CGG		
	CCA		GCA		AGA		
	CCG		GCG		AGG		

Table 135 | Optimal Codons for the Stramenopile *P. multistriata*



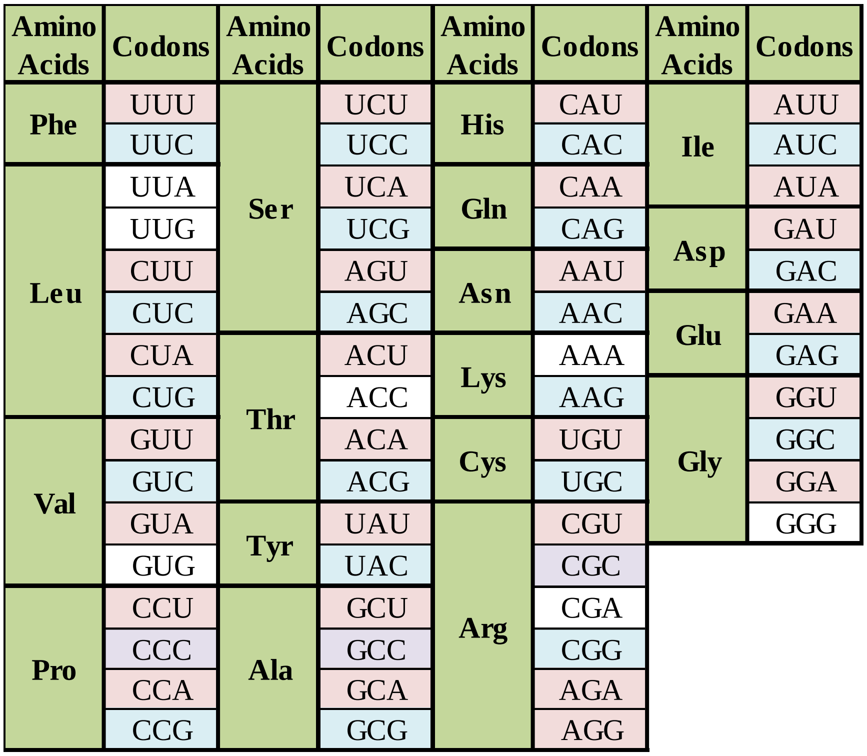
Amino Acids	Codons	Amino Acids	Codons	Amino Acids	Codons	Amino Acids	Codons
Phe	UUU	**Ser**	UCU	**His**	CAU	**Ile**	AUU
	UUC		UCC		CAC		AUC
Leu	UUA		UCA	**Gln**	CAA		AUA
	UUG		UCG		CAG	**Asp**	GAU
	CUU		AGU	**Asn**	AAU		GAC
	CUC		AGC		AAC	**Glu**	GAA
	CUA	**Thr**	ACU	**Lys**	AAA		GAG
	CUG		ACC		AAG	**Gly**	GGU
Val	GUU		ACA	**Cys**	UGU		GGC
	GUC		ACG		UGC		GGA
	GUA	**Tyr**	UAU	**Arg**	CGU		GGG
	GUG		UAC		CGC		
Pro	CCU	**Ala**	GCU		CGA		
	CCC		GCC		CGG		
	CCA		GCA		AGA		
	CCG		GCG		AGG		

Table 136 | Optimal Codons for the Stramenopile *T. pseudonana*



Amino Acids	Codons	Amino Acids	Codons	Amino Acids	Codons	Amino Acids	Codons
Phe	UUU	Ser	UCU	His	CAU	Ile	AUU
	UUC		UCC		CAC		AUC
Leu	UUA		UCA	Gln	CAA		AUA
	UUG		UCG		CAG	Asp	GAU
	CUU		AGU	Asn	AAU		GAC
	CUC		AGC		AAC	Glu	GAA
	CUA	Thr	ACU	Lys	AAA		GAG
	CUG		ACC		AAG	Gly	GGU
Val	GUU		ACA	Cys	UGU		GGC
	GUC		ACG		UGC		GGA
	GUA	Tyr	UAU	Arg	CGU		GGG
	GUG		UAC		CGC		
Pro	CCU	Ala	GCU		CGA		
	CCC		GCC		CGG		
	CCA		GCA		AGA		
	CCG		GCG		AGG		

Table 137 | Optimal Codons for the Stramenopile *N. gaditana*



Amino Acids	Codons	Amino Acids	Codons	Amino Acids	Codons	Amino Acids	Codons
Phe	UUU	Ser	UCU	His	CAU	Ile	AUU
	UUC		UCC		CAC		AUC
Leu	UUA		UCA	Gln	CAA		AUA
	UUG		UCG		CAG	Asp	GAU
	CUU		AGU	Asn	AAU		GAC
	CUC		AGC		AAC	Glu	GAA
	CUA	Thr	ACU	Lys	AAA		GAG
	CUG		ACC		AAG	Gly	GGU
Val	GUU		ACA	Cys	UGU		GGC
	GUC		ACG		UGC		GGA
	GUA	Tyr	UAU	Arg	CGU		GGG
	GUG		UAC		CGC		
Pro	CCU	Ala	GCU		CGA		
	CCC		GCC		CGG		
	CCA		GCA		AGA		
	CCG		GCG		AGG		

Table 138 | Optimal Codons for the Stramenopile *A. anophagefferens*

Amino Acids	Codons	Amino Acids	Codons	Amino Acids	Codons	Amino Acids	Codons
Phe	UUU	**Ser**	UCU	**His**	CAU	**Ile**	AUU
	UUC		UCC		CAC		AUC
Leu	UUA		UCA	**Gln**	CAA		AUA
	UUG		UCG		CAG	**Asp**	GAU
	CUU		AGU	**Asn**	AAU		GAC
	CUC		AGC		AAC	**Glu**	GAA
	CUA	**Thr**	ACU	**Lys**	AAA		GAG
	CUG		ACC		AAG	**Gly**	GGU
Val	GUU		ACA	**Cys**	UGU		GGC
	GUC		ACG		UGC		GGA
	GUA	**Tyr**	UAU	**Arg**	CGU		GGG
	GUG		UAC		CGC		
Pro	CCU	**Ala**	GCU		CGA		
	CCC		GCC		CGG		
	CCA		GCA		AGA		
	CCG		GCG		AGG		

Table 139 | Optimal Codons for the Rhizarian *B. natans*

Amino Acids	Codons	Amino Acids	Codons	Amino Acids	Codons	Amino Acids	Codons
Phe	UUU	**Ser**	UCU	**His**	CAU	**Ile**	AUU
	UUC		UCC		CAC		AUC
Leu	UUA		UCA	**Gln**	CAA		AUA
	UUG		UCG		CAG	**Asp**	GAU
	CUU		AGU	**Asn**	AAU		GAC
	CUC		AGC		AAC	**Glu**	GAA
	CUA	**Thr**	ACU	**Lys**	AAA		GAG
	CUG		ACC		AAG	**Gly**	GGU
Val	GUU		ACA	**Cys**	UGU		GGC
	GUC		ACG		UGC		GGA
	GUA	**Tyr**	UAU	**Arg**	CGU		GGG
	GUG		UAC		CGC		
Pro	CCU	**Ala**	GCU		CGA		
	CCC		GCC		CGG		
	CCA		GCA		AGA		
	CCG		GCG		AGG		

Table 140 | Optimal Codons for the Rhizarian *P. brassicae*

6.6 SeC tRNA genes

Organism	Sequence Name	tRNA Type	Anticodon	Isotype Model
A. castellanii	KB007909	SeC	TCA	SeC
D. discoideum	4	SeC	TCA	SeC
P. fungivorum	scaffold214_0	SeC	TCA	SeC
T. trahens	GL349473	SeC	TCA	SeC
C. owczarzaki	KE346370	SeC	TCA	SeC
S. arctica	KQ242101	SeC	TCA	SeC
L. major	6	SeC	TCA	SeC
L. pyrrhocoris	LpyrH10_03	SeC	TCA	SeC
L. seymouri	Lsey_0079	SeC	TCA	SeC
T. brucei	9	SeC	TCA	SeC
T. cruzi	TcCL_ESM_contig_00012 TcCL_NonESM_contig_00066 TcCL_Unassigned_contig_00513	SeC SeC SeC	TCA TCA TCA	SeC SeC SeC
T. theileri	TM35_Tth_2_V1	SeC	TCA	SeC
T. thermophila	GG662308	SeC	TCA	SeC
S. coeruleus	SteCoe_contig_43	SeC	TCA	SeC
P. gallinaceum	PGAL8A_v1_76	SeC	TCA	SeC
P. falciparum	14	SeC	TCA	SeC
P. vivax	12	SeC	TCA	SeC
P. infestans	supercont1-1	SeC	TCA	SeC
P. sojae	PHYSOscaffold_5	SeC	TCA	SeC
A. astaci	scaffold_4916	SeC	TCA	SeC
A. invadans	scf_2373_170_contig_1	SeC	TCA	SeC
S. parasitica	KK583202	SeC	TCA	SeC
P. multistriata	contig30 contig168 contig246	SeC SeC SeC	TCA TCA TCA	SeC SeC SeC
T. pseudonana	18 9	SeC SeC	TCA TCA	SeC SeC
N. gaditana	16	SeC	TCA	SeC
A. anophagefferens	GL833131	SeC	TCA	SeC
B. natans	scaffold_5	SeC	TCA	SeC

Table 141 | SeC tRNA genes found within the 33 protists

6.7 tRNA genes found for each organism

tRNA type	Anticodon	No. found	tRNA type	Anticodon	No. found	tRNA type	Anticodon	No. found	tRNA type	Anticodon	No. found
Phe	AAA	0	**Ser**	AGA	8	**His**	ATG	0	**Ile**	AAT	11
	GAA	12		GGA	0		GTG	6		GAT	1
Leu	TAA	1		TGA	3	**Gln**	TTG	4		TAT	4
	CAA	2		CGA	6		CTG	16	**Asp**	ATC	1
	AAG	21		ACT	0	**Asn**	ATT	2		GTC	24
	GAG	0		GCT	11		GTT	21	**Glu**	TTC	7
	TAG	4	**Thr**	AGT	20	**Lys**	TTT	4		CTC	22
	CAG	10		GGT	0		CTT	32	**Gly**	ACC	0
Val	AAC	14		TGT	2	**Cys**	ACA	3		GCC	23
	GAC	0		CGT	6		GCA	10		TCC	5
	TAC	2	**Tyr**	ATA	1	**Arg**	ACG	16		CCC	4
	CAC	11		GTA	13		GCG	0	**Met**	CAT	11
Pro	AGG	13	**Ala**	AGC	23		TCG	4	**Trp**	CCA	14
	GGG	0		GGC	0		CCG	3			
	TGG	2		TGC	3		TCT	3	**Total Anticodons found:**		
	CGG	9		CGC	8		CCT	8			464

Table 142 | Number and type of tRNA genes found for the Amoebozoan *A. castellani* (chimeric and pseudo genes includeed)

tRNA type	Anticodon	No. found	tRNA type	Anticodon	No. found	tRNA type	Anticodon	No. found	tRNA type	Anticodon	No. found
Phe	AAA	0	**Ser**	AGA	0	**His**	ATG	0	**Ile**	AAT	1
	GAA	2		GGA	0		GTG	1		GAT	0
Leu	TAA	1		TGA	1	**Gln**	TTG	3		TAT	29
	CAA	1		CGA	0		CTG	1	**Asp**	ATC	0
	AAG	2		ACT	0	**Asn**	ATT	0		GTC	3
	GAG	0		GCT	0		GTT	1	**Glu**	TTC	3
	TAG	12	**Thr**	AGT	3	**Lys**	TTT	0		CTC	1
	CAG	0		GGT	0		CTT	1	**Gly**	ACC	0
Val	AAC	0		TGT	1	**Cys**	ACA	0		GCC	0
	GAC	2		CGT	0		GCA	0		TCC	1
	TAC	0	**Tyr**	ATA	0	**Arg**	ACG	2		CCC	4
	CAC	1		GTA	2		GCG	0	**Met**	CAT	0
Pro	AGG	0	**Ala**	AGC	3		TCG	0	**Trp**	CCA	0
	GGG	0		GGC	0		CCG	1			
	TGG	1		TGC	0		TCT	0	**Total Anticodons found:**		
	CGG	2		CGC	0		CCT	2			88

Table 143 | Number and type of tRNA genes found for the Amoebozoan *E. histolytica* (chimeric and pseudo genes included)

tRNA type	Anticodon	No. found	tRNA type	Anticodon	No. found	tRNA type	Anticodon	No. found	tRNA type	Anticodon	No. found
Phe	AAA	0	Ser	AGA	9	His	ATG	0	Ile	AAT	17
	GAA	15		GGA	0		GTG	9		GAT	0
Leu	TAA	19		TGA	16	Gln	TTG	13		TAT	4
	CAA	4		CGA	1		CTG	1	Asp	ATC	0
	AAG	11		ACT	0	Asn	ATT	0		GTC	22
	GAG	0		GCT	12		GTT	18	Glu	TTC	19
	TAG	3	Thr	AGT	18	Lys	TTT	22		CTC	3
	CAG	1		GGT	0		CTT	10	Gly	ACC	0
Val	AAC	20		TGT	6	Cys	ACA	0		GCC	18
	GAC	0		CGT	1		GCA	8		TCC	5
	TAC	7	Tyr	ATA	0	Arg	ACG	7		CCC	0
	CAC	1		GTA	12		GCG	0	Met	CAT	7
Pro	AGG	1	Ala	AGC	15		TCG	1	Trp	CCA	7
	GGG	0		GGC	0		CCG	0			
	TGG	15		TGC	8		TCT	11	Total Anticodons found:		
	CGG	0		CGC	0		CCT	1			398

Table 144 | Number and type of tRNA genes found for the Amoebozoan *D. discoideum* (chimeric and pseudo genes included)

tRNA type	Anticodon	No. found	tRNA type	Anticodon	No. found	tRNA type	Anticodon	No. found	tRNA type	Anticodon	No. found
Phe	AAA	0	Ser	AGA	12	His	ATG	0	Ile	AAT	15
	GAA	12		GGA	0		GTG	8		GAT	0
Leu	TAA	4		TGA	3	Gln	TTG	6		TAT	2
	CAA	6		CGA	4		CTG	14	Asp	ATC	0
	AAG	11		ACT	0	Asn	ATT	0		GTC	16
	GAG	0		GCT	8		GTT	15	Glu	TTC	6
	TAG	2	Thr	AGT	13	Lys	TTT	6		CTC	16
	CAG	8		GGT	0		CTT	18	Gly	ACC	0
Val	AAC	13		TGT	5	Cys	ACA	0		GCC	14
	GAC	0		CGT	3		GCA	6		TCC	20
	TAC	3	Tyr	ATA	0	Arg	ACG	8		CCC	1
	CAC	4		GTA	14		GCG	0	Met	CAT	7
Pro	AGG	10	Ala	AGC	18		TCG	5	Trp	CCA	10
	GGG	0		GGC	0		CCG	1			
	TGG	5		TGC	7		TCT	10	Total Anticodons found:		
	CGG	2		CGC	2		CCT	3			376

Table 145 | Number and type of tRNA genes found for the Amoebozoan *P. fungivorum* (chimeric and pseudo genes included)

tRNA type	Anticodon	No. found	tRNA type	Anticodon	No. found	tRNA type	Anticodon	No. found	tRNA type	Anticodon	No. found
Phe	AAA	0	Ser	AGA	0	His	ATG	0	Ile	AAT	6
	GAA	4		GGA	0		GTG	0		GAT	0
Leu	TAA	2		TGA	2	Gln	TTG	0		TAT	1
	CAA	3		CGA	4		CTG	0	Asp	ATC	0
	AAG	3		ACT	0	Asn	ATT	0		GTC	3
	GAG	0		GCT	2		GTT	2	Glu	TTC	1
	TAG	0	Thr	AGT	0	Lys	TTT	1		CTC	0
	CAG	2		GGT	0		CTT	1	Gly	ACC	0
Val	AAC	2		TGT	2	Cys	ACA	0		GCC	10
	GAC	0		CGT	2		GCA	0		TCC	2
	TAC	0	Tyr	ATA	0	Arg	ACG	4		CCC	2
	CAC	2		GTA	2		GCG	0	Met	CAT	4
Pro	AGG	0	Ala	AGC	5		TCG	3	Trp	CCA	1
	GGG	0		GGC	0		CCG	1			
	TGG	2		TGC	1		TCT	1	Total Anticodons found:		
	CGG	2		CGC	1		CCT	2			88

Table 146 | Number and type of tRNA genes found for the Apusomonadan *T. trahens* (chimeric and pseudo genes included)

tRNA type	Anticodon	No. found	tRNA type	Anticodon	No. found	tRNA type	Anticodon	No. found	tRNA type	Anticodon	No. found
Phe	AAA	0	Ser	AGA	21	His	ATG	0	Ile	AAT	28
	GAA	14		GGA	0		GTG	18		GAT	1
Leu	TAA	4		TGA	2	Gln	TTG	4		TAT	1
	CAA	1		CGA	6		CTG	17	Asp	ATC	0
	AAG	15		ACT	0	Asn	ATT	0		GTC	31
	GAG	0		GCT	8		GTT	15	Glu	TTC	5
	TAG	2	Thr	AGT	23	Lys	TTT	6		CTC	31
	CAG	11		GGT	0		CTT	38	Gly	ACC	0
Val	AAC	27		TGT	2	Cys	ACA	0		GCC	32
	GAC	0		CGT	2		GCA	11		TCC	1
	TAC	2	Tyr	ATA	1	Arg	ACG	35		CCC	4
	CAC	6		GTA	11		GCG	0	Met	CAT	8
Pro	AGG	16	Ala	AGC	30		TCG	0	Trp	CCA	11
	GGG	0		GGC	0		CCG	3			
	TGG	4		TGC	5		TCT	3	Total Anticodons found:		
	CGG	10		CGC	7		CCT	1			534

Table 147 | Number and type of tRNA genes found for the Opisthokont *F. alba* (chimeric and pseudo genes included)

tRNA type	Anticodon	No. found	tRNA type	Anticodon	No. found	tRNA type	Anticodon	No. found	tRNA type	Anticodon	No. found
Phe	AAA	0	Ser	AGA	1	His	ATG	0	Ile	AAT	3
	GAA	3		GGA	0		GTG	1		GAT	0
Leu	TAA	1		TGA	0	Gln	TTG	1		TAT	0
	CAA	0		CGA	0		CTG	3	Asp	ATC	0
	AAG	4		ACT	0	Asn	ATT	0		GTC	7
	GAG	0		GCT	0		GTT	2	Glu	TTC	2
	TAG	1	Thr	AGT	3	Lys	TTT	1		CTC	5
	CAG	2		GGT	0		CTT	5	Gly	ACC	0
Val	AAC	4		TGT	1	Cys	ACA	0		GCC	5
	GAC	0		CGT	2		GCA	1		TCC	3
	TAC	1	Tyr	ATA	0	Arg	ACG	3		CCC	0
	CAC	1		GTA	2		GCG	0	Met	CAT	4
Pro	AGG	3	Ala	AGC	6		TCG	1	Trp	CCA	1
	GGG	0		GGC	0		CCG	0			
	TGG	0		TGC	2		TCT	1	Total Anticodons found:		
	CGG	2		CGC	2		CCT	1			91

Table 148 | Number and type of tRNA genes found for the Opisthokont *C. owczarzaki* (chimeric and pseudo genes included)

tRNA type	Anticodon	No. found	tRNA type	Anticodon	No. found	tRNA type	Anticodon	No. found	tRNA type	Anticodon	No. found
Phe	AAA	0	Ser	AGA	7	His	ATG	0	Ile	AAT	13
	GAA	11		GGA	0		GTG	8		GAT	1
Leu	TAA	4		TGA	4	Gln	TTG	7		TAT	3
	CAA	5		CGA	3		CTG	7	Asp	ATC	0
	AAG	3		ACT	0	Asn	ATT	0		GTC	15
	GAG	0		GCT	5		GTT	12	Glu	TTC	5
	TAG	2	Thr	AGT	9	Lys	TTT	8		CTC	12
	CAG	1		GGT	1		CTT	15	Gly	ACC	1
Val	AAC	7		TGT	6	Cys	ACA	0		GCC	16
	GAC	1		CGT	4		GCA	7		TCC	11
	TAC	10	Tyr	ATA	0	Arg	ACG	12		CCC	0
	CAC	5		GTA	8		GCG	0	Met	CAT	11
Pro	AGG	10	Ala	AGC	12		TCG	5	Trp	CCA	6
	GGG	0		GGC	0		CCG	0			
	TGG	7		TGC	10		TCT	6	Total Anticodons found:		
	CGG	2		CGC	5		CCT	2			325

Table 149 | Number and type of tRNA genes found for the Opisthokont *S. arctica* (chimeric and pseudo genes included)

tRNA type	Anticodon	No. found	tRNA type	Anticodon	No. found	tRNA type	Anticodon	No. found	tRNA type	Anticodon	No. found
Phe	AAA	0	**Ser**	AGA	1	**His**	ATG	0	**Ile**	AAT	0
	GAA	1		GGA	0		GTG	1		GAT	13
Leu	TAA	1		TGA	3	**Gln**	TTG	3		TAT	1
	CAA	1		CGA	1		CTG	1	**Asp**	ATC	0
	AAG	2		ACT	0	**Asn**	ATT	0		GTC	3
	GAG	0		GCT	1		GTT	2	**Glu**	TTC	1
	TAG	1	**Thr**	AGT	1	**Lys**	TTT	1		CTC	2
	CAG	1		GGT	0		CTT	2	**Gly**	ACC	0
Val	AAC	0		TGT	1	**Cys**	ACA	0		GCC	2
	GAC	2		CGT	1		GCA	1		TCC	1
	TAC	1	**Tyr**	ATA	0	**Arg**	ACG	2		CCC	1
	CAC	1		GTA	1		GCG	0	**Met**	CAT	3
Pro	AGG	1	**Ala**	AGC	2		TCG	1	**Trp**	CCA	1
	GGG	0		GGC	0		CCG	1			
	TGG	1		TGC	1		TCT	1	**Total Anticodons found:**		
	CGG	1		CGC	1		CCT	1			73

Table 150 | Number and type of tRNA genes found for the Excavate *G. lamblia* (chimeric and pseudo genes included)

tRNA type	Anticodon	No. found	tRNA type	Anticodon	No. found	tRNA type	Anticodon	No. found	tRNA type	Anticodon	No. found
Phe	AAA	0	**Ser**	AGA	5	**His**	ATG	0	**Ile**	AAT	0
	GAA	6		GGA	0		GTG	3		GAT	6
Leu	TAA	4		TGA	1	**Glu**	TTC	5		TAT	2
	CAA	2		CGA	1		CTC	3	**Asp**	ATC	0
	AAG	5		ACT	0	**Asn**	ATT	1		GTC	6
	GAG	0		GCT	5		GTT	9	**Gln**	TTG	1
	TAG	1	**Thr**	AGT	5	**Lys**	TTT	2		CTG	2
	CAG	1		GGT	0		CTT	5		TTA	-
Val	AAC	0		TGT	1	**Cys**	ACA	0		CTA	-
	GAC	3		CGT	2		GCA	4	**Gly**	ACC	0
	TAC	2	**Tyr**	ATA	0	**Arg**	ACG	4		GCC	0
	CAC	1		GTA	3		GCG	0		TCC	2
Pro	AGG	2	**Ala**	AGC	7		TCG	1		CCC	1
	GGG	0		GGC	0		CCG	1	**Met**	CAT	5
	TGG	2		TGC	2		TCT	3	**Trp**	CCA	2
	CGG	1		CGC	1		CCT	2	Total Anticodons found:		133

Table 151 | Number and type of tRNA genes found for the Excavate *S. salmonicida* (chimeric and pseudo genes included)

tRNA type	Anticodon	No. found	tRNA type	Anticodon	No. found	tRNA type	Anticodon	No. found	tRNA type	Anticodon	No. found
Phe	AAA	0	**Ser**	AGA	1	**His**	ATG	0	**Ile**	AAT	3
	GAA	2		GGA	0		GTG	2		GAT	0
Leu	TAA	1		TGA	1	**Gln**	TTG	1		TAT	1
	CAA	1		CGA	1		CTG	3	**Asp**	ATC	0
	AAG	3		ACT	0	**Asn**	ATT	0		GTC	3
	GAG	0		GCT	2		GTT	3	**Glu**	TTC	1
	TAG	1	**Thr**	AGT	3	**Lys**	TTT	1		CTC	2
	CAG	2		GGT	0		CTT	4	**Gly**	ACC	0
Val	AAC	2		TGT	1	**Cys**	ACA	0		GCC	4
	GAC	0		CGT	2		GCA	1		TCC	1
	TAC	1	**Tyr**	ATA	0	**Arg**	ACG	4		CCC	1
	CAC	2		GTA	3		GCG	0	**Met**	CAT	2
Pro	AGG	2	**Ala**	AGC	2		TCG	1	**Trp**	CCA	1
	GGG	0		GGC	0		CCG	1			
	TGG	1		TGC	1		TCT	1	**Total Anticodons found:**		
	CGG	2		CGC	2		CCT	1			81

Table 152 | Number and type of tRNA genes found for the Excavate *L. major* (chimeric and pseudo genes included)

tRNA type	Anticodon	No. found	tRNA type	Anticodon	No. found	tRNA type	Anticodon	No. found	tRNA type	Anticodon	No. found
Phe	AAA	0	**Ser**	AGA	6	**His**	ATG	0	**Ile**	AAT	4
	GAA	2		GGA	0		GTG	2		GAT	0
Leu	TAA	1		TGA	0	**Gln**	TTG	1		TAT	1
	CAA	1		CGA	0		CTG	3	**Asp**	ATC	0
	AAG	3		ACT	0	**Asn**	ATT	0		GTC	4
	GAG	0		GCT	0		GTT	3	**Glu**	TTC	2
	TAG	1	**Thr**	AGT	3	**Lys**	TTT	1		CTC	4
	CAG	3		GGT	0		CTT	5	**Gly**	ACC	0
Val	AAC	3		TGT	1	**Cys**	ACA	0		GCC	4
	GAC	0		CGT	3		GCA	2		TCC	1
	TAC	1	**Tyr**	ATA	0	**Arg**	ACG	7		CCC	1
	CAC	4		GTA	3		GCG	0	**Met**	CAT	3
Pro	AGG	2	**Ala**	AGC	2		TCG	1	**Trp**	CCA	2
	GGG	0		GGC	0		CCG	1			
	TGG	1		TGC	1		TCT	1	**Total Anticodons found:**		
	CGG	3		CGC	3		CCT	1			101

Table 153 | Number and type of tRNA genes found for the Excavate *L. pyrrhocoris* (chimeric and pseudo genes included)

tRNA type	Anticodon	No. found	tRNA type	Anticodon	No. found	tRNA type	Anticodon	No. found	tRNA type	Anticodon	No. found
Phe	AAA	0	**Ser**	AGA	1	**His**	ATG	0	**Ile**	AAT	4
	GAA	2		GGA	0		GTG	2		GAT	0
Leu	TAA	1		TGA	1	**Gln**	TTG	1		TAT	1
	CAA	1		CGA	1		CTG	2	**Asp**	ATC	0
	AAG	2		ACT	0	**Asn**	ATT	0		GTC	3
	GAG	0		GCT	2		GTT	3	**Glu**	TTC	2
	TAG	1	**Thr**	AGT	3	**Lys**	TTT	1		CTC	3
	CAG	3		GGT	0		CTT	3	**Gly**	ACC	0
Val	AAC	3		TGT	1	**Cys**	ACA	0		GCC	4
	GAC	0		CGT	2		GCA	2		TCC	1
	TAC	1	**Tyr**	ATA	0	**Arg**	ACG	6		CCC	1
	CAC	4		GTA	3		GCG	0	**Met**	CAT	3
Pro	AGG	2	**Ala**	AGC	3		TCG	1	**Trp**	CCA	2
	GGG	0		GGC	0		CCG	1			
	TGG	1		TGC	1		TCT	1	**Total Anticodons found:**		
	CGG	3		CGC	2		CCT	1			92

Table 154 | Number and type of tRNA genes found for the Excavate *L. seymouri* (chimeric and pseudo genes included)

tRNA type	Anticodon	No. found	tRNA type	Anticodon	No. found	tRNA type	Anticodon	No. found	tRNA type	Anticodon	No. found
Phe	AAA	0	**Ser**	AGA	1	**His**	ATG	0	**Ile**	AAT	2
	GAA	2		GGA	0		GTG	1		GAT	0
Leu	TAA	1		TGA	1	**Gln**	TTG	1		TAT	1
	CAA	1		CGA	1		CTG	2	**Asp**	ATC	0
	AAG	1		ACT	0	**Asn**	ATT	0		GTC	2
	GAG	0		GCT	2		GTT	2	**Glu**	TTC	1
	TAG	1	**Thr**	AGT	1	**Lys**	TTT	1		CTC	2
	CAG	1		GGT	0		CTT	3	**Gly**	ACC	0
Val	AAC	2		TGT	1	**Cys**	ACA	0		GCC	3
	GAC	0		CGT	1		GCA	2		TCC	1
	TAC	1	**Tyr**	ATA	0	**Arg**	ACG	3		CCC	1
	CAC	1		GTA	1		GCG	0	**Met**	CAT	2
Pro	AGG	1	**Ala**	AGC	2		TCG	1	**Trp**	CCA	1
	GGG	0		GGC	0		CCG	1			
	TGG	1		TGC	1		TCT	1	**Total Anticodons found:**		
	CGG	1		CGC	2		CCT	1			63

Table 155 | Number and type of tRNA genes found for the Excavate *T. brucei* (chimeric and pseudo genes included)

tRNA type	Anticodon	No. found	tRNA type	Anticodon	No. found	tRNA type	Anticodon	No. found	tRNA type	Anticodon	No. found
Phe	AAA	0	**Ser**	AGA	1	**His**	ATG	0	**Ile**	AAT	4
	GAA	3		GGA	0		GTG	1		GAT	0
Leu	TAA	2		TGA	2	**Gln**	TTG	2		TAT	4
	CAA	1		CGA	2		CTG	4	**Asp**	ATC	0
	AAG	3		ACT	0	**Asn**	ATT	0		GTC	2
	GAG	0		GCT	2		GTT	3	**Glu**	TTC	2
	TAG	2	**Thr**	AGT	2	**Lys**	TTT	2		CTC	4
	CAG	2		GGT	0		CTT	4	**Gly**	ACC	0
Val	AAC	2		TGT	2	**Cys**	ACA	0		GCC	2
	GAC	0		CGT	1		GCA	2		TCC	2
	TAC	2	**Tyr**	ATA	1	**Arg**	ACG	3		CCC	2
	CAC	2		GTA	2		GCG	0	**Met**	CAT	4
Pro	AGG	1	**Ala**	AGC	1		TCG	2	**Trp**	CCA	2
	GGG	0		GGC	0		CCG	2			
	TGG	1		TGC	2		TCT	2	**Total Anticodons found:**		
	CGG	1		CGC	2		CCT	2			99

Table 156 | Number and type of tRNA genes found for the Excavate *T. cruzi* (chimeric and pseudo genes included)

tRNA type	Anticodon	No. found	tRNA type	Anticodon	No. found	tRNA type	Anticodon	No. found	tRNA type	Anticodon	No. found
Phe	AAA	0	**Ser**	AGA	3	**His**	ATG	0	**Ile**	AAT	5
	GAA	14		GGA	0		GTG	2		GAT	0
Leu	TAA	3		TGA	2	**Gln**	TTG	2		TAT	1
	CAA	2		CGA	1		CTG	4	**Asp**	ATC	0
	AAG	3		ACT	0	**Asn**	ATT	0		GTC	1
	GAG	0		GCT	2		GTT	5	**Glu**	TTC	4
	TAG	1	**Thr**	AGT	3	**Lys**	TTT	2		CTC	7
	CAG	5		GGT	0		CTT	7	**Gly**	ACC	0
Val	AAC	3		TGT	3	**Cys**	ACA	0		GCC	4
	GAC	0		CGT	1		GCA	2		TCC	5
	TAC	4	**Tyr**	ATA	0	**Arg**	ACG	7		CCC	1
	CAC	5		GTA	4		GCG	0	**Met**	CAT	4
Pro	AGG	3	**Ala**	AGC	2		TCG	2	**Trp**	CCA	2
	GGG	0		GGC	0		CCG	1			
	TGG	2		TGC	3		TCT	1	**Total Anticodons found:**		
	CGG	2		CGC	4		CCT	1			145

Table 157 | Number and type of tRNA genes found for the Excavate *T. theileri* (chimeric and pseudo genes included)

tRNA type	Anticodon	No. found	tRNA type	Anticodon	No. found	tRNA type	Anticodon	No. found	tRNA type	Anticodon	No. found
Phe	AAA	0	**Ser**	AGA	20	**His**	ATG	0	**Ile**	AAT	29
	GAA	24		GGA	0		GTG	16		GAT	0
Leu	TAA	20		TGA	9	**Glu**	TTC	36		TAT	13
	CAA	14		CGA	2		CTC	7	**Asp**	ATC	0
	AAG	18		ACT	0	**Asn**	ATT	0		GTC	31
	GAG	0		GCT	15		GTT	31	**Gln**	TTG	37
	TAG	3	**Thr**	AGT	23	**Lys**	TTT	29		CTG	7
	CAG	2		GGT	0		CTT	33		TTA	-
Val	AAC	21		TGT	6	**Cys**	ACA	0		CTA	-
	GAC	0		CGT	1		GCA	18	**Gly**	ACC	0
	TAC	6	**Tyr**	ATA	0	**Arg**	ACG	8		GCC	31
	CAC	3		GTA	19		GCG	0		TCC	12
Pro	AGG	20	**Ala**	AGC	32		TCG	1		CCC	1
	GGG	0		GGC	0		CCG	1	**Met**	CAT	12
	TGG	7		TGC	11		TCT	25	**Trp**	CCA	21
	CGG	1		CGC	2		CCT	2	Total Anticodons found:		680

Table 158 | Number and type of tRNA genes found for the Alveolate *T. thermophila* (chimeric and pseudo genes included)

tRNA type	Anticodon	No. found	tRNA type	Anticodon	No. found	tRNA type	Anticodon	No. found	tRNA type	Anticodon	No. found
Phe	AAA	0	**Ser**	AGA	6	**His**	ATG	0	**Ile**	AAT	9
	GAA	5		GGA	0		GTG	6		GAT	1
Leu	TAA	6		TGA	5	**Gln**	TTG	6		TAT	9
	CAA	7		CGA	3		CTG	5	**Asp**	ATC	10
	AAG	7		ACT	0	**Asn**	ATT	0		GTC	2
	GAG	0		GCT	7		GTT	9	**Glu**	TTC	9
	TAG	2	**Thr**	AGT	8	**Lys**	TTT	11		CTC	4
	CAG	2		GGT	0		CTT	5	**Gly**	ACC	0
Val	AAC	4		TGT	4	**Cys**	ACA	0		GCC	8
	GAC	0		CGT	4		GCA	7		TCC	5
	TAC	4	**Tyr**	ATA	0	**Arg**	ACG	3		CCC	3
	CAC	3		GTA	12		GCG	0	**Met**	CAT	9
Pro	AGG	5	**Ala**	AGC	10		TCG	2	**Trp**	CCA	6
	GGG	0		GGC	0		CCG	2			
	TGG	6		TGC	3		TCT	7	**Total Anticodons found:**		
	CGG	3		CGC	2		CCT	3			259

Table 159 | Number and type of tRNA genes found for the Alveolate *S. coeruleus* (chimeric and pseudo genes included)

tRNA type	Anticodon	No. found	tRNA type	Anticodon	No. found	tRNA type	Anticodon	No. found	tRNA type	Anticodon	No. found
Phe	AAA	0	**Ser**	AGA	1	**His**	ATG	0	**Ile**	AAT	1
	GAA	1		GGA	0		GTG	1		GAT	0
Leu	TAA	1		TGA	1	**Gln**	TTG	1		TAT	1
	CAA	1		CGA	1		CTG	1	**Asp**	ATC	0
	AAG	1		ACT	0	**Asn**	ATT	0		GTC	1
	GAG	0		GCT	1		GTT	1	**Glu**	TTC	1
	TAG	1	**Thr**	AGT	1	**Lys**	TTT	1		CTC	1
	CAG	0		GGT	0		CTT	2	**Gly**	ACC	0
Val	AAC	1		TGT	1	**Cys**	ACA	0		GCC	1
	GAC	0		CGT	1		GCA	1		TCC	1
	TAC	1	**Tyr**	ATA	0	**Arg**	ACG	1		CCC	0
	CAC	1		GTA	1		GCG	0	**Met**	CAT	1
Pro	AGG	1	**Ala**	AGC	1		TCG	1	**Trp**	CCA	1
	GGG	0		GGC	0		CCG	0			
	TGG	1		TGC	1		TCT	1	**Total Anticodons found:**		
	CGG	1		CGC	1		CCT	1			43

Table 160 | Number and type of tRNA genes found for the Alveolate *P. gallinaceum* (chimeric and pseudo genes included)

tRNA type	Anticodon	No. found	tRNA type	Anticodon	No. found	tRNA type	Anticodon	No. found	tRNA type	Anticodon	No. found
Phe	AAA	0	**Ser**	AGA	1	**His**	ATG	0	**Ile**	AAT	1
	GAA	1		GGA	0		GTG	1		GAT	0
Leu	TAA	1		TGA	1	**Gln**	TTG	1		TAT	1
	CAA	1		CGA	1		CTG	1	**Asp**	ATC	0
	AAG	1		ACT	0	**Asn**	ATT	0		GTC	1
	GAG	0		GCT	1		GTT	1	**Glu**	TTC	1
	TAG	1	**Thr**	AGT	1	**Lys**	TTT	1		CTC	1
	CAG	1		GGT	0		CTT	1	**Gly**	ACC	0
Val	AAC	1		TGT	1	**Cys**	ACA	0		GCC	1
	GAC	0		CGT	1		GCA	1		TCC	1
	TAC	1	**Tyr**	ATA	0	**Arg**	ACG	1		CCC	0
	CAC	1		GTA	1		GCG	0	**Met**	CAT	1
Pro	AGG	1	**Ala**	AGC	1		TCG	1	**Trp**	CCA	1
	GGG	0		GGC	0		CCG	0			
	TGG	1		TGC	1		TCT	1	**Total Anticodons found:**		
	CGG	1		CGC	1		CCT	1			43

Table 161 | Number and type of tRNA genes found for the Alveolate *P. falciparum* (chimeric and pseudo genes included)

tRNA type	Anticodon	No. found	tRNA type	Anticodon	No. found	tRNA type	Anticodon	No. found	tRNA type	Anticodon	No. found
Phe	AAA	0	**Ser**	AGA	1	**His**	ATG	0	**Ile**	AAT	1
	GAA	2		GGA	0		GTG	1		GAT	0
Leu	TAA	1		TGA	1	**Gln**	TTG	1		TAT	1
	CAA	1		CGA	1		CTG	1	**Asp**	ATC	0
	AAG	1		ACT	0	**Asn**	ATT	0		GTC	1
	GAG	0		GCT	1		GTT	1	**Glu**	TTC	1
	TAG	1	**Thr**	AGT	1	**Lys**	TTT	1		CTC	1
	CAG	1		GGT	0		CTT	1	**Gly**	ACC	0
Val	AAC	1		TGT	1	**Cys**	ACA	0		GCC	1
	GAC	0		CGT	1		GCA	1		TCC	1
	TAC	1	**Tyr**	ATA	0	**Arg**	ACG	1		CCC	0
	CAC	1		GTA	1		GCG	0	**Met**	CAT	1
Pro	AGG	1	**Ala**	AGC	1		TCG	1	**Trp**	CCA	1
	GGG	0		GGC	0		CCG	0			
	TGG	1		TGC	1		TCT	1	**Total Anticodons found:**		
	CGG	1		CGC	1		CCT	1			44

Table 162 | Number and type of tRNA genes found for the Alveolate *P. vivax* (chimeric and pseudo genes included)

tRNA type	Anticodon	No. found	tRNA type	Anticodon	No. found	tRNA type	Anticodon	No. found	tRNA type	Anticodon	No. found
Phe	AAA	3	**Ser**	AGA	129	**His**	ATG	2	**Ile**	AAT	132
	GAA	151		GGA	25		GTG	112		GAT	8
Leu	TAA	192		TGA	548	**Gln**	TTG	214		TAT	19
	CAA	202		CGA	79		CTG	231	**Asp**	ATC	6
	AAG	87		ACT	0	**Asn**	ATT	3		GTC	531
	GAG	0		GCT	122		GTT	161	**Glu**	TTC	172
	TAG	24	**Thr**	AGT	159	**Lys**	TTT	128		CTC	353
	CAG	61		GGT	0		CTT	127	**Gly**	ACC	6
Val	AAC	130		TGT	80	**Cys**	ACA	0		GCC	75
	GAC	3		CGT	155		GCA	104		TCC	50
	TAC	56	**Tyr**	ATA	0	**Arg**	ACG	79		CCC	60
	CAC	342		GTA	91		GCG	5	**Met**	CAT	172
Pro	AGG	103	**Ala**	AGC	280		TCG	21	**Trp**	CCA	120
	GGG	0		GGC	17		CCG	4			
	TGG	533		TGC	97		TCT	19	**Total Anticodons found:**		
	CGG	132		CGC	134		CCT	3			6852

Table 163 | Number and type of tRNA genes found for the Stramenopile *P. infestans* (chimeric and pseudo genes included)

tRNA type	Anticodon	No. found	tRNA type	Anticodon	No. found	tRNA type	Anticodon	No. found	tRNA type	Anticodon	No. found
Phe	AAA	0	**Ser**	AGA	60	**His**	ATG	0	**Ile**	AAT	31
	GAA	113		GGA	0		GTG	60		GAT	0
Leu	TAA	8		TGA	3	**Gln**	TTG	20		TAT	0
	CAA	16		CGA	32		CTG	22	**Asp**	ATC	0
	AAG	29		ACT	0	**Asn**	ATT	0		GTC	5
	GAG	0		GCT	7		GTT	5	**Glu**	TTC	73
	TAG	4	**Thr**	AGT	28	**Lys**	TTT	8		CTC	172
	CAG	10		GGT	0		CTT	133	**Gly**	ACC	0
Val	AAC	61		TGT	7	**Cys**	ACA	0		GCC	42
	GAC	0		CGT	28		GCA	7		TCC	38
	TAC	8	**Tyr**	ATA	0	**Arg**	ACG	40		CCC	38
	CAC	38		GTA	8		GCG	0	**Met**	CAT	50
Pro	AGG	70	**Ala**	AGC	71		TCG	10	**Trp**	CCA	8
	GGG	0		GGC	0		CCG	6			
	TGG	10		TGC	33		TCT	3	**Total Anticodons found:**		
	CGG	6		CGC	216		CCT	50			1687

Table 164 | Number and type of tRNA genes found for the Stramenopile *P. sojae* (chimeric and pseudo genes included)

tRNA type	Anticodon	No. found	tRNA type	Anticodon	No. found	tRNA type	Anticodon	No. found	tRNA type	Anticodon	No. found
Phe	AAA	0	**Ser**	AGA	2	**His**	ATG	0	**Ile**	AAT	5
	GAA	1		GGA	0		GTG	2		GAT	3
Leu	TAA	1		TGA	4	**Gln**	TTG	3		TAT	4
	CAA	2		CGA	1		CTG	3	**Asp**	ATC	1
	AAG	2		ACT	0	**Asn**	ATT	1		GTC	1
	GAG	0		GCT	5		GTT	8	**Glu**	TTC	5
	TAG	2	**Thr**	AGT	4	**Lys**	TTT	4		CTC	6
	CAG	1		GGT	0		CTT	1	**Gly**	ACC	0
Val	AAC	1		TGT	4	**Cys**	ACA	1		GCC	4
	GAC	0		CGT	2		GCA	15		TCC	6
	TAC	4	**Tyr**	ATA	0	**Arg**	ACG	2		CCC	1
	CAC	4		GTA	2		GCG	1	**Met**	CAT	8
Pro	AGG	2	**Ala**	AGC	2		TCG	1	**Trp**	CCA	1
	GGG	0		GGC	0		CCG	2			
	TGG	3		TGC	8		TCT	4	**Total Anticodons found:**		
	CGG	1		CGC	1		CCT	2			154

Table 165 | Number and type of tRNA genes found for the Stramenopile *P. halstedii* (chimeric and pseudo genes included)

tRNA type	Anticodon	No. found	tRNA type	Anticodon	No. found	tRNA type	Anticodon	No. found	tRNA type	Anticodon	No. found
Phe	AAA	0	**Ser**	AGA	6	**His**	ATG	0	**Ile**	AAT	2
	GAA	3		GGA	1		GTG	3		GAT	0
Leu	TAA	2		TGA	2	**Gln**	TTG	4		TAT	3
	CAA	18		CGA	7		CTG	8	**Asp**	ATC	0
	AAG	2		ACT	0	**Asn**	ATT	0		GTC	11
	GAG	0		GCT	11		GTT	4	**Glu**	TTC	4
	TAG	2	**Thr**	AGT	4	**Lys**	TTT	9		CTC	3
	CAG	9		GGT	2		CTT	100	**Gly**	ACC	0
Val	AAC	0		TGT	5	**Cys**	ACA	0		GCC	3
	GAC	7		CGT	2		GCA	12		TCC	13
	TAC	5	**Tyr**	ATA	0	**Arg**	ACG	5		CCC	5
	CAC	6		GTA	3		GCG	0	**Met**	CAT	3
Pro	AGG	12	**Ala**	AGC	3		TCG	3	**Trp**	CCA	10
	GGG	0		GGC	0		CCG	2			
	TGG	6		TGC	3		TCT	6	**Total Anticodons found:**		
	CGG	5		CGC	5		CCT	2			346

Table 166 | Number and type of tRNA genes found for the Stramenopile *A. astaci* (chimeric and pseudo genes included)

tRNA type	Anticodon	No. found	tRNA type	Anticodon	No. found	tRNA type	Anticodon	No. found	tRNA type	Anticodon	No. found
Phe	AAA	0	**Ser**	AGA	7	**His**	ATG	0	**Ile**	AAT	4
	GAA	5		GGA	0		GTG	1		GAT	1
Leu	TAA	2		TGA	4	**Gln**	TTG	8		TAT	1
	CAA	22		CGA	23		CTG	7	**Asp**	ATC	0
	AAG	10		ACT	0	**Asn**	ATT	0		GTC	15
	GAG	1		GCT	12		GTT	6	**Glu**	TTC	2
	TAG	4	**Thr**	AGT	3	**Lys**	TTT	8		CTC	3
	CAG	10		GGT	0		CTT	14	**Gly**	ACC	0
Val	AAC	0		TGT	28	**Cys**	ACA	0		GCC	12
	GAC	11		CGT	7		GCA	11		TCC	11
	TAC	7	**Tyr**	ATA	0	**Arg**	ACG	9		CCC	6
	CAC	5		GTA	2		GCG	0	**Met**	CAT	4
Pro	AGG	2	**Ala**	AGC	5		TCG	4	**Trp**	CCA	10
	GGG	0		GGC	1		CCG	1			
	TGG	8		TGC	5		TCT	5	**Total Anticodons found:**		
	CGG	5		CGC	5		CCT	2			339

Table 167 | Number and type of tRNA genes found for the Stramenopile *A. invadans* (chimeric and pseudo genes included)

tRNA type	Anticodon	No. found	tRNA type	Anticodon	No. found	tRNA type	Anticodon	No. found	tRNA type	Anticodon	No. found
Phe	AAA	2	**Ser**	AGA	10	**His**	ATG	0	**Ile**	AAT	2
	GAA	6		GGA	0		GTG	3		GAT	1
Leu	TAA	3		TGA	3	**Gln**	TTG	5		TAT	3
	CAA	8		CGA	5		CTG	1	**Asp**	ATC	0
	AAG	5		ACT	0	**Asn**	ATT	0		GTC	7
	GAG	0		GCT	7		GTT	9	**Glu**	TTC	2
	TAG	4	**Thr**	AGT	8	**Lys**	TTT	4		CTC	1
	CAG	5		GGT	0		CTT	11	**Gly**	ACC	0
Val	AAC	1		TGT	3	**Cys**	ACA	0		GCC	2
	GAC	5		CGT	5		GCA	5		TCC	7
	TAC	16	**Tyr**	ATA	0	**Arg**	ACG	0		CCC	5
	CAC	17		GTA	5		GCG	1	**Met**	CAT	13
Pro	AGG	5	**Ala**	AGC	6		TCG	4	**Trp**	CCA	4
	GGG	0		GGC	0		CCG	4			
	TGG	4		TGC	7		TCT	3	**Total Anticodons found:**		
	CGG	3		CGC	0		CCT	2			242

Table 168 | Number and type of tRNA genes found for the Stramenopile *S. parasitica* (chimeric and pseudo genes included)

tRNA type	Anticodon	No. found	tRNA type	Anticodon	No. found	tRNA type	Anticodon	No. found	tRNA type	Anticodon	No. found
Phe	AAA	0	**Ser**	AGA	7	**His**	ATG	0	**Ile**	AAT	8
	GAA	8		GGA	0		GTG	4		GAT	0
Leu	TAA	1		TGA	2	**Gln**	TTG	5		TAT	0
	CAA	5		CGA	4		CTG	3	**Asp**	ATC	0
	AAG	9		ACT	0	**Asn**	ATT	0		GTC	12
	GAG	0		GCT	4		GTT	8	**Glu**	TTC	7
	TAG	4	**Thr**	AGT	8	**Lys**	TTT	6		CTC	4
	CAG	1		GGT	0		CTT	8	**Gly**	ACC	0
Val	AAC	7		TGT	3	**Cys**	ACA	0		GCC	7
	GAC	0		CGT	3		GCA	0		TCC	11
	TAC	3	**Tyr**	ATA	0	**Arg**	ACG	0		CCC	0
	CAC	1		GTA	5		GCG	0	**Met**	CAT	6
Pro	AGG	5	**Ala**	AGC	12		TCG	4	**Trp**	CCA	4
	GGG	0		GGC	0		CCG	0			
	TGG	5		TGC	5		TCT	3	**Total Anticodons found:**		
	CGG	1		CGC	2		CCT	2			207

Table 169 | Number and type of tRNA genes found for the Stramenopile *P. multistriata* (chimeric and pseudo genes included)

tRNA type	Anticodon	No. found	tRNA type	Anticodon	No. found	tRNA type	Anticodon	No. found	tRNA type	Anticodon	No. found
Phe	AAA	0	Ser	AGA	2	His	ATG	0	Ile	AAT	3
	GAA	3		GGA	0		GTG	2		GAT	0
Leu	TAA	1		TGA	2	Gln	TTG	1		TAT	0
	CAA	2		CGA	1		CTG	2	Asp	ATC	0
	AAG	1		ACT	0	Asn	ATT	0		GTC	4
	GAG	0		GCT	3		GTT	4	Glu	TTC	3
	TAG	1	Thr	AGT	2	Lys	TTT	3		CTC	3
	CAG	1		GGT	0		CTT	3	Gly	ACC	0
Val	AAC	2		TGT	2	Cys	ACA	0		GCC	3
	GAC	0		CGT	1		GCA	0		TCC	4
	TAC	2	Tyr	ATA	0	Arg	ACG	0		CCC	0
	CAC	1		GTA	3		GCG	0	Met	CAT	5
Pro	AGG	2	Ala	AGC	4		TCG	1	Trp	CCA	2
	GGG	0		GGC	0		CCG	1			
	TGG	1		TGC	1		TCT	2	Total Anticodons found:		
	CGG	1		CGC	1		CCT	1			87

Table 170 | Number and type of tRNA genes found for the Stramenopile *T. pseudonana* (chimeric and pseudo genes included)

tRNA type	Anticodon	No. found	tRNA type	Anticodon	No. found	tRNA type	Anticodon	No. found	tRNA type	Anticodon	No. found
Phe	AAA	0	Ser	AGA	1	His	ATG	0	Ile	AAT	1
	GAA	3		GGA	0		GTG	2		GAT	3
Leu	TAA	3		TGA	2	Gln	TTG	3		TAT	0
	CAA	2		CGA	1		CTG	1	Asp	ATC	0
	AAG	1		ACT	0	Asn	ATT	0		GTC	3
	GAG	0		GCT	3		GTT	2	Glu	TTC	3
	TAG	2	Thr	AGT	0	Lys	TTT	3		CTC	1
	CAG	1		GGT	0		CTT	1	Gly	ACC	0
Val	AAC	0		TGT	2	Cys	ACA	0		GCC	3
	GAC	1		CGT	1		GCA	2		TCC	2
	TAC	3	Tyr	ATA	0	Arg	ACG	2		CCC	1
	CAC	1		GTA	3		GCG	0	Met	CAT	7
Pro	AGG	0	Ala	AGC	1		TCG	2	Trp	CCA	2
	GGG	0		GGC	0		CCG	0			
	TGG	3		TGC	4		TCT	2	Total Anticodons found:		
	CGG	1		CGC	1		CCT	1			87

Table 171 | Number and type of tRNA genes found for the Stramenopile *N. gaditana* (chimeric and pseudo genes included)

tRNA type	Anticodon	No. found	tRNA type	Anticodon	No. found	tRNA type	Anticodon	No. found	tRNA type	Anticodon	No. found
Phe	AAA	0	**Ser**	AGA	1	**His**	ATG	0	**Ile**	AAT	2
	GAA	2		GGA	0		GTG	2		GAT	0
Leu	TAA	1		TGA	1	**Gln**	TTG	1		TAT	3
	CAA	1		CGA	2		CTG	0	**Asp**	ATC	0
	AAG	3		ACT	0	**Asn**	ATT	0		GTC	4
	GAG	0		GCT	1		GTT	2	**Glu**	TTC	1
	TAG	1	**Thr**	AGT	1	**Lys**	TTT	1		CTC	2
	CAG	1		GGT	0		CTT	3	**Gly**	ACC	0
Val	AAC	0		TGT	1	**Cys**	ACA	0		GCC	4
	GAC	1		CGT	2		GCA	1		TCC	1
	TAC	1	**Tyr**	ATA	0	**Arg**	ACG	4		CCC	1
	CAC	0		GTA	0		GCG	0	**Met**	CAT	1
Pro	AGG	3	**Ala**	AGC	0		TCG	2	**Trp**	CCA	1
	GGG	0		GGC	0		CCG	2			
	TGG	1		TGC	1		TCT	2	**Total Anticodons found:**		
	CGG	1		CGC	2		CCT	1			69

Table 172 | Number and type of tRNA genes found for the Stramenopile *A. anophagefferens* (chimeric and pseudo genes included)

tRNA type	Anticodon	No. found	tRNA type	Anticodon	No. found	tRNA type	Anticodon	No. found	tRNA type	Anticodon	No. found
Phe	AAA	0	**Ser**	AGA	1	**His**	ATG	0	**Ile**	AAT	3
	GAA	6		GGA	0		GTG	1		GAT	0
Leu	TAA	2		TGA	2	**Gln**	TTG	2		TAT	1
	CAA	2		CGA	3		CTG	2	**Asp**	ATC	0
	AAG	4		ACT	1	**Asn**	ATT	0		GTC	7
	GAG	0		GCT	1		GTT	3	**Glu**	TTC	6
	TAG	1	**Thr**	AGT	2	**Lys**	TTT	3		CTC	2
	CAG	2		GGT	0		CTT	3	**Gly**	ACC	0
Val	AAC	2		TGT	2	**Cys**	ACA	3		GCC	2
	GAC	0		CGT	2		GCA	1		TCC	3
	TAC	2	**Tyr**	ATA	0	**Arg**	ACG	2		CCC	3
	CAC	2		GTA	3		GCG	0	**Met**	CAT	3
Pro	AGG	4	**Ala**	AGC	1		TCG	1	**Trp**	CCA	2
	GGG	0		GGC	0		CCG	1			
	TGG	4		TGC	4		TCT	1	**Total Anticodons found:**		
	CGG	2		CGC	0		CCT	1			109

Table 173 | Number and type of tRNA genes found for the Rhizarian *B. natans* (chimeric and pseudo genes included)

tRNA type	Anticodon	No. found	tRNA type	Anticodon	No. found	tRNA type	Anticodon	No. found	tRNA type	Anticodon	No. found
Phe	AAA	0	**Ser**	AGA	3	**His**	ATG	0	**Ile**	AAT	6
	GAA	4		GGA	0		GTG	3		GAT	0
Leu	TAA	1		TGA	1	**Gln**	TTG	2		TAT	1
	CAA	2		CGA	4		CTG	4	**Asp**	ATC	0
	AAG	3		ACT	0	**Asn**	ATT	0		GTC	8
	GAG	0		GCT	2		GTT	12	**Glu**	TTC	4
	TAG	1	**Thr**	AGT	3	**Lys**	TTT	2		CTC	5
	CAG	5		GGT	0		CTT	7	**Gly**	ACC	0
Val	AAC	5		TGT	2	**Cys**	ACA	0		GCC	7
	GAC	0		CGT	3		GCA	3		TCC	2
	TAC	1	**Tyr**	ATA	0	**Arg**	ACG	5		CCC	2
	CAC	3		GTA	3		GCG	0	**Met**	CAT	5
Pro	AGG	3	**Ala**	AGC	5		TCG	3	**Trp**	CCA	2
	GGG	0		GGC	0		CCG	2			
	TGG	2		TGC	3		TCT	1	**Total Anticodons found:**		
	CGG	4		CGC	4		CCT	1			154

Table 174 | Number and type of tRNA genes found for the Rhizarian *P. brassicae* (chimeric and pseudo genes included)

tRNA type	Anticodon	No. found	tRNA type	Anticodon	No. found	tRNA type	Anticodon	No. found	tRNA type	Anticodon	No. found
Phe	AAA	0	**Ser**	AGA	8	**His**	ATG	0	**Ile**	AAT	10
	GAA	0		GGA	0		GTG	6		GAT	0
Leu	TAA	1		TGA	2	**Gln**	TTG	4		TAT	2
	CAA	2		CGA	5		CTG	16	**Asp**	ATC	0
	AAG	21		ACT	0	**Asn**	ATT	0		GTC	0
	GAG	0		GCT	11		GTT	0	**Glu**	TTC	5
	TAG	3	**Thr**	AGT	19	**Lys**	TTT	4		CTC	21
	CAG	10		GGT	0		CTT	31	**Gly**	ACC	0
Val	AAC	12		TGT	2	**Cys**	ACA	0		GCC	21
	GAC	0		CGT	6		GCA	9		TCC	5
	TAC	2	**Tyr**	ATA	0	**Arg**	ACG	16		CCC	4
	CAC	10		GTA	13		GCG	0	**Met**	CAT	9
Pro	AGG	11	**Ala**	AGC	23		TCG	4	**Trp**	CCA	13
	GGG	0		GGC	0		CCG	3			
	TGG	2		TGC	3		TCT	2	**Total Anticodons found:**		
	CGG	7		CGC	8		CCT	8			374

Table 175 | Number and type of tRNA genes found for the Amoebozoan *A. castellani* (chimeric and pseudo genes removed)

tRNA type	Anticodon	No. found	tRNA type	Anticodon	No. found	tRNA type	Anticodon	No. found	tRNA type	Anticodon	No. found
Phe	AAA	0	**Ser**	AGA	0	**His**	ATG	0	**Ile**	AAT	1
	GAA	0		GGA	0		GTG	1		GAT	0
Leu	TAA	0		TGA	1	**Gln**	TTG	1		TAT	17
	CAA	1		CGA	0		CTG	1	**Asp**	ATC	0
	AAG	2		ACT	0	**Asn**	ATT	0		GTC	3
	GAG	0		GCT	0		GTT	0	**Glu**	TTC	0
	TAG	12	**Thr**	AGT	3	**Lys**	TTT	0		CTC	0
	CAG	0		GGT	0		CTT	1	**Gly**	ACC	0
Val	AAC	0		TGT	1	**Cys**	ACA	0		GCC	1
	GAC	2		CGT	0		GCA	0		TCC	1
	TAC	0	**Tyr**	ATA	0	**Arg**	ACG	2		CCC	3
	CAC	0		GTA	1		GCG	0	**Met**	CAT	0
Pro	AGG	0	**Ala**	AGC	1		TCG	0	**Trp**	CCA	0
	GGG	0		GGC	0		CCG	1			
	TGG	0		TGC	0		TCT	0	**Total Anticodons found:**		
	CGG	2		CGC	0		CCT	2			61

Table 176 | Number and type of tRNA genes found for the Amoebozoan *E. histolytica* (chimeric and pseudo genes removed)

tRNA type	Anticodon	No. found	tRNA type	Anticodon	No. found	tRNA type	Anticodon	No. found	tRNA type	Anticodon	No. found
Phe	AAA	0	**Ser**	AGA	9	**His**	ATG	0	**Ile**	AAT	17
	GAA	15		GGA	0		GTG	9		GAT	0
Leu	TAA	18		TGA	15	**Gln**	TTG	13		TAT	4
	CAA	4		CGA	1		CTG	1	**Asp**	ATC	0
	AAG	11		ACT	0	**Asn**	ATT	0		GTC	22
	GAG	0		GCT	12		GTT	0	**Glu**	TTC	0
	TAG	3	**Thr**	AGT	18	**Lys**	TTT	22		CTC	1
	CAG	0		GGT	0		CTT	10	**Gly**	ACC	0
Val	AAC	20		TGT	6	**Cys**	ACA	0		GCC	18
	GAC	0		CGT	0		GCA	8		TCC	5
	TAC	7	**Tyr**	ATA	0	**Arg**	ACG	0		CCC	0
	CAC	1		GTA	0		GCG	0	**Met**	CAT	7
Pro	AGG	1	**Ala**	AGC	15		TCG	0	**Trp**	CCA	0
	GGG	0		GGC	0		CCG	0			
	TGG	15		TGC	8		TCT	11	**Total Anticodons found:**		
	CGG	0		CGC	0		CCT	0			327

Table 177 | Number and type of tRNA genes found for the Amoebozoan *D. discoideum* (chimeric and pseudo genes removed)

tRNA type	Anticodon	No. found	tRNA type	Anticodon	No. found	tRNA type	Anticodon	No. found	tRNA type	Anticodon	No. found
Phe	AAA	0	**Ser**	AGA	12	**His**	ATG	0	**Ile**	AAT	15
	GAA	12		GGA	0		GTG	8		GAT	0
Leu	TAA	3		TGA	2	**Gln**	TTG	5		TAT	2
	CAA	5		CGA	4		CTG	13	**Asp**	ATC	0
	AAG	11		ACT	0	**Asn**	ATT	0		GTC	16
	GAG	0		GCT	6		GTT	13	**Glu**	TTC	5
	TAG	2	**Thr**	AGT	13	**Lys**	TTT	5		CTC	16
	CAG	7		GGT	0		CTT	17	**Gly**	ACC	0
Val	AAC	13		TGT	5	**Cys**	ACA	0		GCC	14
	GAC	0		CGT	3		GCA	6		TCC	20
	TAC	3	**Tyr**	ATA	0	**Arg**	ACG	8		CCC	1
	CAC	4		GTA	14		GCG	0	**Met**	CAT	6
Pro	AGG	10	**Ala**	AGC	18		TCG	0	**Trp**	CCA	6
	GGG	0		GGC	0		CCG	1			
	TGG	4		TGC	6		TCT	0	**Total Anticodons found:**		
	CGG	2		CGC	2		CCT	0			338

Table 178 | Number and type of tRNA genes found for the Amoebozoan *P. fungivorum* (chimeric and pseudo genes removed)

tRNA type	Anticodon	No. found	tRNA type	Anticodon	No. found	tRNA type	Anticodon	No. found	tRNA type	Anticodon	No. found
Phe	AAA	0	**Ser**	AGA	0	**His**	ATG	0	**Ile**	AAT	6
	GAA	0		GGA	0		GTG	0		GAT	0
Leu	TAA	2		TGA	2	**Gln**	TTG	0		TAT	0
	CAA	3		CGA	4		CTG	0	**Asp**	ATC	0
	AAG	3		ACT	0	**Asn**	ATT	0		GTC	3
	GAG	0		GCT	2		GTT	2	**Glu**	TTC	0
	TAG	0	**Thr**	AGT	0	**Lys**	TTT	0		CTC	0
	CAG	2		GGT	0		CTT	1	**Gly**	ACC	0
Val	AAC	2		TGT	2	**Cys**	ACA	0		GCC	10
	GAC	0		CGT	2		GCA	0		TCC	2
	TAC	0	**Tyr**	ATA	0	**Arg**	ACG	4		CCC	2
	CAC	1		GTA	2		GCG	0	**Met**	CAT	0
Pro	AGG	0	**Ala**	AGC	5		TCG	0	**Trp**	CCA	0
	GGG	0		GGC	0		CCG	1			
	TGG	2		TGC	1		TCT	0	**Total Anticodons found:**		
	CGG	2		CGC	1		CCT	0			69

Table 179 | Number and type of tRNA genes found for the Apusomonadan *T. trahens* (chimeric and pseudo genes removed)

tRNA type	Anticodon	No. found	tRNA type	Anticodon	No. found	tRNA type	Anticodon	No. found	tRNA type	Anticodon	No. found
Phe	AAA	0	**Ser**	AGA	21	**His**	ATG	0	**Ile**	AAT	27
	GAA	13		GGA	0		GTG	15		GAT	0
Leu	TAA	1		TGA	2	**Gln**	TTG	1		TAT	1
	CAA	1		CGA	6		CTG	17	**Asp**	ATC	0
	AAG	15		ACT	0	**Asn**	ATT	0		GTC	30
	GAG	0		GCT	8		GTT	1	**Glu**	TTC	4
	TAG	1	**Thr**	AGT	22	**Lys**	TTT	3		CTC	31
	CAG	11		GGT	0		CTT	38	**Gly**	ACC	0
Val	AAC	27		TGT	0	**Cys**	ACA	0		GCC	32
	GAC	0		CGT	2		GCA	10		TCC	0
	TAC	1	**Tyr**	ATA	0	**Arg**	ACG	34		CCC	0
	CAC	6		GTA	11		GCG	0	**Met**	CAT	2
Pro	AGG	16	**Ala**	AGC	30		TCG	0	**Trp**	CCA	0
	GGG	0		GGC	0		CCG	3			
	TGG	2		TGC	3		TCT	2	**Total Anticodons found:**		
	CGG	10		CGC	7		CCT	1			468

Table 180 | Number and type of tRNA genes found for the Opisthokont *F. alba* (chimeric and pseudo genes removed)

tRNA type	Anticodon	No. found	tRNA type	Anticodon	No. found	tRNA type	Anticodon	No. found	tRNA type	Anticodon	No. found
Phe	AAA	0	**Ser**	AGA	1	**His**	ATG	0	**Ile**	AAT	3
	GAA	0		GGA	0		GTG	1		GAT	0
Leu	TAA	0		TGA	0	**Gln**	TTG	0		TAT	0
	CAA	0		CGA	0		CTG	3	**Asp**	ATC	0
	AAG	4		ACT	0	**Asn**	ATT	0		GTC	0
	GAG	0		GCT	0		GTT	0	**Glu**	TTC	0
	TAG	0	**Thr**	AGT	3	**Lys**	TTT	1		CTC	0
	CAG	2		GGT	0		CTT	5	**Gly**	ACC	0
Val	AAC	0		TGT	0	**Cys**	ACA	0		GCC	5
	GAC	0		CGT	2		GCA	1		TCC	2
	TAC	1	**Tyr**	ATA	0	**Arg**	ACG	3		CCC	0
	CAC	0		GTA	2		GCG	0	**Met**	CAT	0
Pro	AGG	3	**Ala**	AGC	6		TCG	0	**Trp**	CCA	0
	GGG	0		GGC	0		CCG	0			
	TGG	0		TGC	1		TCT	0	**Total Anticodons found:**		
	CGG	2		CGC	2		CCT	1			54

Table 181 | Number and type of tRNA genes found for the Opisthokont *C. owczarzaki* (chimeric and pseudo genes removed)

tRNA type	Anticodon	No. found	tRNA type	Anticodon	No. found	tRNA type	Anticodon	No. found	tRNA type	Anticodon	No. found
Phe	AAA	0	**Ser**	AGA	6	**His**	ATG	0	**Ile**	AAT	13
	GAA	9		GGA	0		GTG	0		GAT	0
Leu	TAA	3		TGA	4	**Gln**	TTG	0		TAT	0
	CAA	5		CGA	3		CTG	5	**Asp**	ATC	0
	AAG	3		ACT	0	**Asn**	ATT	0		GTC	15
	GAG	0		GCT	5		GTT	0	**Glu**	TTC	0
	TAG	2	**Thr**	AGT	0	**Lys**	TTT	6		CTC	2
	CAG	1		GGT	0		CTT	15	**Gly**	ACC	0
Val	AAC	7		TGT	0	**Cys**	ACA	0		GCC	16
	GAC	0		CGT	4		GCA	6		TCC	10
	TAC	10	**Tyr**	ATA	0	**Arg**	ACG	1		CCC	0
	CAC	5		GTA	8		GCG	0	**Met**	CAT	0
Pro	AGG	0	**Ala**	AGC	2		TCG	0	**Trp**	CCA	0
	GGG	0		GGC	0		CCG	0			
	TGG	5		TGC	4		TCT	0	**Total Anticodons found:**		
	CGG	2		CGC	5		CCT	0			182

Table 182 | Number and type of tRNA genes found for the Opisthokont *S. arctica* (chimeric and pseudo genes removed)

tRNA type	Anticodon	No. found	tRNA type	Anticodon	No. found	tRNA type	Anticodon	No. found	tRNA type	Anticodon	No. found
Phe	AAA	0	**Ser**	AGA	1	**His**	ATG	0	**Ile**	AAT	0
	GAA	1		GGA	0		GTG	1		GAT	0
Leu	TAA	1		TGA	1	**Gln**	TTG	1		TAT	1
	CAA	1		CGA	1		CTG	1	**Asp**	ATC	0
	AAG	2		ACT	0	**Asn**	ATT	0		GTC	2
	GAG	0		GCT	1		GTT	2	**Glu**	TTC	1
	TAG	1	**Thr**	AGT	1	**Lys**	TTT	1		CTC	2
	CAG	1		GGT	0		CTT	2	**Gly**	ACC	0
Val	AAC	0		TGT	1	**Cys**	ACA	0		GCC	2
	GAC	2		CGT	1		GCA	1		TCC	1
	TAC	0	**Tyr**	ATA	0	**Arg**	ACG	2		CCC	1
	CAC	1		GTA	1		GCG	0	**Met**	CAT	2
Pro	AGG	1	**Ala**	AGC	2		TCG	1	**Trp**	CCA	0
	GGG	0		GGC	0		CCG	1			
	TGG	1		TGC	1		TCT	1	**Total Anticodons found:**		
	CGG	1		CGC	1		CCT	1			52

Table 183 | Number and type of tRNA genes found for the Excavate *G. lamblia* (chimeric and pseudo genes removed)

tRNA type	Anticodon	No. found	tRNA type	Anticodon	No. found	tRNA type	Anticodon	No. found	tRNA type	Anticodon	No. found
Phe	AAA	0	**Ser**	AGA	5	**His**	ATG	0	**Ile**	AAT	0
	GAA	5		GGA	0		GTG	3		GAT	0
Leu	TAA	3		TGA	1	**Glu**	TTC	5		TAT	2
	CAA	2		CGA	1		CTC	0	**Asp**	ATC	0
	AAG	5		ACT	0	**Asn**	ATT	0		GTC	6
	GAG	0		GCT	5		GTT	0	**Gln**	TTG	1
	TAG	1	**Thr**	AGT	4	**Lys**	TTT	0		CTG	2
	CAG	1		GGT	0		CTT	5		TTA	-
Val	AAC	0		TGT	1	**Cys**	ACA	0		CTA	-
	GAC	3		CGT	1		GCA	4	**Gly**	ACC	0
	TAC	2	**Tyr**	ATA	0	**Arg**	ACG	4		GCC	1
	CAC	1		GTA	3		GCG	0		TCC	2
Pro	AGG	2	**Ala**	AGC	7		TCG	1		CCC	1
	GGG	0		GGC	0		CCG	1	**Met**	CAT	0
	TGG	2		TGC	2		TCT	3	**Trp**	CCA	0
	CGG	1		CGC	1		CCT	2	Total Anticodons found:		102

Table 184 | Number and type of tRNA genes found for the Excavate *S. salmonicida* (chimeric and pseudo genes removed)

tRNA type	Anticodon	No. found	tRNA type	Anticodon	No. found	tRNA type	Anticodon	No. found	tRNA type	Anticodon	No. found
Phe	AAA	0	**Ser**	AGA	1	**His**	ATG	0	**Ile**	AAT	3
	GAA	2		GGA	0		GTG	2		GAT	0
Leu	TAA	1		TGA	1	**Gln**	TTG	1		TAT	1
	CAA	1		CGA	1		CTG	3	**Asp**	ATC	0
	AAG	3		ACT	0	**Asn**	ATT	0		GTC	3
	GAG	0		GCT	2		GTT	3	**Glu**	TTC	1
	TAG	1	**Thr**	AGT	3	**Lys**	TTT	1		CTC	2
	CAG	2		GGT	0		CTT	3	**Gly**	ACC	0
Val	AAC	2		TGT	1	**Cys**	ACA	0		GCC	4
	GAC	0		CGT	2		GCA	1		TCC	1
	TAC	1	**Tyr**	ATA	0	**Arg**	ACG	4		CCC	1
	CAC	0		GTA	3		GCG	0	**Met**	CAT	2
Pro	AGG	2	**Ala**	AGC	2		TCG	1	**Trp**	CCA	0
	GGG	0		GGC	0		CCG	1			
	TGG	1		TGC	1		TCT	1	**Total Anticodons found:**		
	CGG	2		CGC	2		CCT	1			77

Table 185 | Number and type of tRNA genes found for the Excavate *L. major* (chimeric and pseudo genes removed)

tRNA type	Anticodon	No. found	tRNA type	Anticodon	No. found	tRNA type	Anticodon	No. found	tRNA type	Anticodon	No. found
Phe	AAA	0	**Ser**	AGA	6	**His**	ATG	0	**Ile**	AAT	4
	GAA	2		GGA	0		GTG	2		GAT	0
Leu	TAA	1		TGA	0	**Gln**	TTG	1		TAT	1
	CAA	1		CGA	0		CTG	3	**Asp**	ATC	0
	AAG	3		ACT	0	**Asn**	ATT	0		GTC	4
	GAG	0		GCT	0		GTT	3	**Glu**	TTC	2
	TAG	1	**Thr**	AGT	3	**Lys**	TTT	1		CTC	4
	CAG	3		GGT	0		CTT	5	**Gly**	ACC	0
Val	AAC	3		TGT	1	**Cys**	ACA	0		GCC	4
	GAC	0		CGT	3		GCA	2		TCC	1
	TAC	1	**Tyr**	ATA	0	**Arg**	ACG	7		CCC	1
	CAC	0		GTA	3		GCG	0	**Met**	CAT	3
Pro	AGG	2	**Ala**	AGC	2		TCG	1	**Trp**	CCA	0
	GGG	0		GGC	0		CCG	1			
	TGG	1		TGC	1		TCT	1	**Total Anticodons found:**		
	CGG	3		CGC	3		CCT	1			95

Table 186 | Number and type of tRNA genes found for the Excavate *L. pyrrhocoris* (chimeric and pseudo genes removed)

tRNA type	Anticodon	No. found	tRNA type	Anticodon	No. found	tRNA type	Anticodon	No. found	tRNA type	Anticodon	No. found
Phe	AAA	0	**Ser**	AGA	1	**His**	ATG	0	**Ile**	AAT	4
	GAA	2		GGA	0		GTG	2		GAT	0
Leu	TAA	1		TGA	1	**Gln**	TTG	1		TAT	1
	CAA	1		CGA	1		CTG	2	**Asp**	ATC	0
	AAG	2		ACT	0	**Asn**	ATT	0		GTC	3
	GAG	0		GCT	2		GTT	3	**Glu**	TTC	2
	TAG	1	**Thr**	AGT	3	**Lys**	TTT	1		CTC	3
	CAG	3		GGT	0		CTT	3	**Gly**	ACC	0
Val	AAC	3		TGT	1	**Cys**	ACA	0		GCC	4
	GAC	0		CGT	2		GCA	2		TCC	1
	TAC	1	**Tyr**	ATA	0	**Arg**	ACG	6		CCC	1
	CAC	0		GTA	3		GCG	0	**Met**	CAT	3
Pro	AGG	2	**Ala**	AGC	3		TCG	1	**Trp**	CCA	0
	GGG	0		GGC	0		CCG	1			
	TGG	1		TGC	1		TCT	1	**Total Anticodons found:**		
	CGG	3		CGC	2		CCT	1			86

Table 187 | Number and type of tRNA genes found for the Excavate *L. seymouri* (chimeric and pseudo genes removed)

tRNA type	Anticodon	No. found	tRNA type	Anticodon	No. found	tRNA type	Anticodon	No. found	tRNA type	Anticodon	No. found
Phe	AAA	0	**Ser**	AGA	1	**His**	ATG	0	**Ile**	AAT	2
	GAA	2		GGA	0		GTG	1		GAT	0
Leu	TAA	1		TGA	1	**Gln**	TTG	1		TAT	1
	CAA	1		CGA	1		CTG	2	**Asp**	ATC	0
	AAG	1		ACT	0	**Asn**	ATT	0		GTC	2
	GAG	0		GCT	2		GTT	2	**Glu**	TTC	1
	TAG	1	**Thr**	AGT	1	**Lys**	TTT	1		CTC	2
	CAG	1		GGT	0		CTT	3	**Gly**	ACC	0
Val	AAC	2		TGT	1	**Cys**	ACA	0		GCC	3
	GAC	0		CGT	1		GCA	2		TCC	1
	TAC	1	**Tyr**	ATA	0	**Arg**	ACG	3		CCC	1
	CAC	0		GTA	1		GCG	0	**Met**	CAT	2
Pro	AGG	1	**Ala**	AGC	2		TCG	1	**Trp**	CCA	0
	GGG	0		GGC	0		CCG	1			
	TGG	1		TGC	1		TCT	1	**Total Anticodons found:**		
	CGG	1		CGC	2		CCT	1			61

Table 188 | Number and type of tRNA genes found for the Excavate *T. brucei* (chimeric and pseudo genes removed)

tRNA type	Anticodon	No. found	tRNA type	Anticodon	No. found	tRNA type	Anticodon	No. found	tRNA type	Anticodon	No. found
Phe	AAA	0	**Ser**	AGA	1	**His**	ATG	0	**Ile**	AAT	4
	GAA	2		GGA	0		GTG	1		GAT	0
Leu	TAA	2		TGA	2	**Gln**	TTG	2		TAT	2
	CAA	1		CGA	2		CTG	4	**Asp**	ATC	0
	AAG	3		ACT	0	**Asn**	ATT	0		GTC	2
	GAG	0		GCT	2		GTT	3	**Glu**	TTC	2
	TAG	2	**Thr**	AGT	2	**Lys**	TTT	2		CTC	4
	CAG	2		GGT	0		CTT	4	**Gly**	ACC	0
Val	AAC	2		TGT	2	**Cys**	ACA	0		GCC	2
	GAC	0		CGT	1		GCA	2		TCC	2
	TAC	2	**Tyr**	ATA	1	**Arg**	ACG	3		CCC	2
	CAC	0		GTA	2		GCG	0	**Met**	CAT	4
Pro	AGG	1	**Ala**	AGC	1		TCG	2	**Trp**	CCA	0
	GGG	0		GGC	0		CCG	2			
	TGG	1		TGC	2		TCT	2	**Total Anticodons found:**		
	CGG	1		CGC	2		CCT	2			92

Table 189 | Number and type of tRNA genes found for the Excavate *T. cruzi* (chimeric and pseudo genes removed)

tRNA type	Anticodon	No. found	tRNA type	Anticodon	No. found	tRNA type	Anticodon	No. found	tRNA type	Anticodon	No. found
Phe	AAA	0	**Ser**	AGA	3	**His**	ATG	0	**Ile**	AAT	5
	GAA	14		GGA	0		GTG	2		GAT	0
Leu	TAA	3		TGA	2	**Gln**	TTG	2		TAT	1
	CAA	2		CGA	1		CTG	4	**Asp**	ATC	0
	AAG	3		ACT	0	**Asn**	ATT	0		GTC	1
	GAG	0		GCT	2		GTT	5	**Glu**	TTC	4
	TAG	1	**Thr**	AGT	3	**Lys**	TTT	2		CTC	7
	CAG	5		GGT	0		CTT	7	**Gly**	ACC	0
Val	AAC	3		TGT	3	**Cys**	ACA	0		GCC	4
	GAC	0		CGT	1		GCA	2		TCC	5
	TAC	4	**Tyr**	ATA	0	**Arg**	ACG	7		CCC	1
	CAC	0		GTA	4		GCG	0	**Met**	CAT	4
Pro	AGG	3	**Ala**	AGC	2		TCG	2	**Trp**	CCA	0
	GGG	0		GGC	0		CCG	1			
	TGG	2		TGC	3		TCT	1	**Total Anticodons found:**		
	CGG	2		CGC	4		CCT	1			138

Table 190 | Number and type of tRNA genes found for the Excavate *T. theileri* (chimeric and pseudo genes removed)

tRNA type	Anticodon	No. found	tRNA type	Anticodon	No. found	tRNA type	Anticodon	No. found	tRNA type	Anticodon	No. found
Phe	AAA	0	**Ser**	AGA	19	**His**	ATG	0	**Ile**	AAT	28
	GAA	24		GGA	0		GTG	0		GAT	0
Leu	TAA	18		TGA	9	**Glu**	TTC	36		TAT	8
	CAA	14		CGA	2		CTC	7	**Asp**	ATC	0
	AAG	18		ACT	0	**Asn**	ATT	0		GTC	31
	GAG	0		GCT	15		GTT	0	**Gln**	TTG	13
	TAG	3	**Thr**	AGT	22	**Lys**	TTT	29		CTC	2
	CAG	1		GGT	0		CTT	33		TTA	-
Val	AAC	21		TGT	6	**Cys**	ACA	0		CTA	-
	GAC	0		CGT	1		GCA	15	**Gly**	ACC	0
	TAC	6	**Tyr**	ATA	0	**Arg**	ACG	6		GCC	31
	CAC	3		GTA	18		GCG	0		TCC	12
Pro	AGG	20	**Ala**	AGC	32		TCG	1		CCC	1
	GGG	0		GGC	0		CCG	1	**Met**	CAT	12
	TGG	7		TGC	11		TCT	24	**Trp**	CCA	12
	CGG	0		CGC	2		CCT	2	Total Anticodons found:		576

Table 191 | Number and type of tRNA genes found for the Alveolate *T. thermophila* (chimeric and pseudo genes removed)

tRNA type	Anticodon	No. found	tRNA type	Anticodon	No. found	tRNA type	Anticodon	No. found	tRNA type	Anticodon	No. found
Phe	AAA	0	**Ser**	AGA	6	**His**	ATG	0	**Ile**	AAT	9
	GAA	5		GGA	0		GTG	5		GAT	1
Leu	TAA	5		TGA	5	**Gln**	TTG	6		TAT	1
	CAA	7		CGA	3		CTG	5	**Asp**	ATC	10
	AAG	7		ACT	0	**Asn**	ATT	0		GTC	0
	GAG	0		GCT	6		GTT	0	**Glu**	TTC	9
	TAG	2	**Thr**	AGT	8	**Lys**	TTT	11		CTC	4
	CAG	2		GGT	0		CTT	5	**Gly**	ACC	0
Val	AAC	4		TGT	4	**Cys**	ACA	0		GCC	8
	GAC	0		CGT	4		GCA	5		TCC	5
	TAC	4	**Tyr**	ATA	0	**Arg**	ACG	0		CCC	3
	CAC	3		GTA	11		GCG	0	**Met**	CAT	8
Pro	AGG	5	**Ala**	AGC	9		TCG	2	**Trp**	CCA	5
	GGG	0		GGC	0		CCG	0			
	TGG	6		TGC	3		TCT	6	**Total Anticodons found:**		
	CGG	3		CGC	2		CCT	3			225

Table 192 | Number and type of tRNA genes found for the Alveolate *S. coeruleus* (chimeric and pseudo genes removed)

tRNA type	Anticodon	No. found	tRNA type	Anticodon	No. found	tRNA type	Anticodon	No. found	tRNA type	Anticodon	No. found
Phe	AAA	0	**Ser**	AGA	1	**His**	ATG	0	**Ile**	AAT	0
	GAA	1		GGA	0		GTG	1		GAT	0
Leu	TAA	1		TGA	1	**Gln**	TTG	1		TAT	1
	CAA	1		CGA	1		CTG	1	**Asp**	ATC	0
	AAG	1		ACT	0	**Asn**	ATT	0		GTC	1
	GAG	0		GCT	1		GTT	1	**Glu**	TTC	1
	TAG	1	**Thr**	AGT	1	**Lys**	TTT	1		CTC	1
	CAG	0		GGT	0		CTT	2	**Gly**	ACC	0
Val	AAC	1		TGT	1	**Cys**	ACA	0		GCC	1
	GAC	0		CGT	1		GCA	1		TCC	1
	TAC	1	**Tyr**	ATA	0	**Arg**	ACG	0		CCC	0
	CAC	1		GTA	1		GCG	0	**Met**	CAT	1
Pro	AGG	1	**Ala**	AGC	1		TCG	1	**Trp**	CCA	1
	GGG	0		GGC	0		CCG	0			
	TGG	1		TGC	1		TCT	1	**Total Anticodons found:**		
	CGG	1		CGC	1		CCT	1			41

Table 193 | Number and type of tRNA genes found for the Alveolate *P. gallinaceum* (chimeric and pseudo genes removed)

tRNA type	Anticodon	No. found	tRNA type	Anticodon	No. found	tRNA type	Anticodon	No. found	tRNA type	Anticodon	No. found
Phe	AAA	0	**Ser**	AGA	1	**His**	ATG	0	**Ile**	AAT	0
	GAA	1		GGA	0		GTG	1		GAT	0
Leu	TAA	1		TGA	1	**Gln**	TTG	1		TAT	1
	CAA	1		CGA	1		CTG	1	**Asp**	ATC	0
	AAG	1		ACT	0	**Asn**	ATT	0		GTC	1
	GAG	0		GCT	1		GTT	1	**Glu**	TTC	1
	TAG	1	**Thr**	AGT	1	**Lys**	TTT	1		CTC	1
	CAG	1		GGT	0		CTT	1	**Gly**	ACC	0
Val	AAC	1		TGT	1	**Cys**	ACA	0		GCC	1
	GAC	0		CGT	1		GCA	1		TCC	1
	TAC	1	**Tyr**	ATA	0	**Arg**	ACG	1		CCC	0
	CAC	1		GTA	1		GCG	0	**Met**	CAT	1
Pro	AGG	1	**Ala**	AGC	1		TCG	1	**Trp**	CCA	1
	GGG	0		GGC	0		CCG	0			
	TGG	1		TGC	1		TCT	1	**Total Anticodons found:**		
	CGG	1		CGC	1		CCT	1			42

Table 194 | Number and type of tRNA genes found for the Alveolate *P. falciparum* (chimeric and pseudo genes removed)

tRNA type	Anticodon	No. found	tRNA type	Anticodon	No. found	tRNA type	Anticodon	No. found	tRNA type	Anticodon	No. found
Phe	AAA	0	**Ser**	AGA	1	**His**	ATG	0	**Ile**	AAT	0
	GAA	1		GGA	0		GTG	1		GAT	0
Leu	TAA	1		TGA	1	**Gln**	TTG	1		TAT	1
	CAA	1		CGA	1		CTG	1	**Asp**	ATC	0
	AAG	1		ACT	0	**Asn**	ATT	0		GTC	1
	GAG	0		GCT	1		GTT	1	**Glu**	TTC	1
	TAG	1	**Thr**	AGT	1	**Lys**	TTT	1		CTC	1
	CAG	1		GGT	0		CTT	1	**Gly**	ACC	0
Val	AAC	1		TGT	1	**Cys**	ACA	0		GCC	1
	GAC	0		CGT	1		GCA	1		TCC	1
	TAC	1	**Tyr**	ATA	0	**Arg**	ACG	1		CCC	0
	CAC	1		GTA	1		GCG	0	**Met**	CAT	1
Pro	AGG	1	**Ala**	AGC	1		TCG	1	**Trp**	CCA	1
	GGG	0		GGC	0		CCG	0			
	TGG	1		TGC	1		TCT	1	**Total Anticodons found:**		
	CGG	1		CGC	1		CCT	1			42

Table 195 | Number and type of tRNA genes found for the Alveolate *P. vivax* (chimeric and pseudo genes removed)

tRNA type	Anticodon	No. found	tRNA type	Anticodon	No. found	tRNA type	Anticodon	No. found	tRNA type	Anticodon	No. found
Phe	AAA	0	**Ser**	AGA	125	**His**	ATG	0	**Ile**	AAT	122
	GAA	143		GGA	0		GTG	105		GAT	1
Leu	TAA	178		TGA	8	**Gln**	TTG	203		TAT	2
	CAA	192		CGA	63		CTG	211	**Asp**	ATC	2
	AAG	83		ACT	0	**Asn**	ATT	0		GTC	197
	GAG	0		GCT	121		GTT	137	**Glu**	TTC	164
	TAG	16	**Thr**	AGT	152	**Lys**	TTT	94		CTC	335
	CAG	58		GGT	0		CTT	109	**Gly**	ACC	1
Val	AAC	25		TGT	63	**Cys**	ACA	0		GCC	69
	GAC	0		CGT	125		GCA	83		TCC	45
	TAC	12	**Tyr**	ATA	0	**Arg**	ACG	78		CCC	58
	CAC	17		GTA	81		GCG	5	**Met**	CAT	150
Pro	AGG	95	**Ala**	AGC	276		TCG	18	**Trp**	CCA	7
	GGG	0		GGC	1		CCG	4			
	TGG	84		TGC	96		TCT	2	**Total Anticodons found:**		
	CGG	127		CGC	126		CCT	2			4471

Table 196 | Number and type of tRNA genes found for the Stramenopile *P. infestans* (chimeric and pseudo genes removed)

tRNA type	Anticodon	No. found	tRNA type	Anticodon	No. found	tRNA type	Anticodon	No. found	tRNA type	Anticodon	No. found
Phe	AAA	0	**Ser**	AGA	60	**His**	ATG	0	**Ile**	AAT	23
	GAA	110		GGA	0		GTG	60		GAT	0
Leu	TAA	7		TGA	3	**Gln**	TTG	18		TAT	4
	CAA	15		CGA	30		CTG	22	**Asp**	ATC	0
	AAG	29		ACT	0	**Asn**	ATT	0		GTC	5
	GAG	0		GCT	7		GTT	5	**Glu**	TTC	55
	TAG	3	**Thr**	AGT	28	**Lys**	TTT	8		CTC	171
	CAG	10		GGT	0		CTT	131	**Gly**	ACC	0
Val	AAC	0		TGT	6	**Cys**	ACA	0		GCC	41
	GAC	0		CGT	28		GCA	7		TCC	38
	TAC	4	**Tyr**	ATA	0	**Arg**	ACG	33		CCC	38
	CAC	8		GTA	8		GCG	0	**Met**	CAT	38
Pro	AGG	70	**Ala**	AGC	71		TCG	10	**Trp**	CCA	0
	GGG	0		GGC	0		CCG	6			
	TGG	10		TGC	33		TCT	3	**Total Anticodons found:**		
	CGG	6		CGC	199		CCT	4			1465

Table 197 | Number and type of tRNA genes found for the Stramenopile *P. sojae* (chimeric and pseudo genes removed)

tRNA type	Anticodon	No. found	tRNA type	Anticodon	No. found	tRNA type	Anticodon	No. found	tRNA type	Anticodon	No. found
Phe	AAA	0	**Ser**	AGA	2	**His**	ATG	0	**Ile**	AAT	5
	GAA	0		GGA	0		GTG	0		GAT	1
Leu	TAA	1		TGA	3	**Gln**	TTG	0		TAT	3
	CAA	1		CGA	0		CTG	3	**Asp**	ATC	0
	AAG	1		ACT	0	**Asn**	ATT	0		GTC	0
	GAG	0		GCT	2		GTT	1	**Glu**	TTC	2
	TAG	1	**Thr**	AGT	3	**Lys**	TTT	2		CTC	2
	CAG	0		GGT	0		CTT	1	**Gly**	ACC	0
Val	AAC	1		TGT	1	**Cys**	ACA	0		GCC	1
	GAC	0		CGT	2		GCA	8		TCC	2
	TAC	0	**Tyr**	ATA	0	**Arg**	ACG	0		CCC	0
	CAC	0		GTA	1		GCG	0	**Met**	CAT	3
Pro	AGG	2	**Ala**	AGC	2		TCG	0	**Trp**	CCA	0
	GGG	0		GGC	0		CCG	2			
	TGG	0		TGC	4		TCT	1	**Total Anticodons found:**		
	CGG	0		CGC	1		CCT	2			67

Table 198 | Number and type of tRNA genes found for the Stramenopile *P. halstedii* (chimeric and pseudo genes removed)

tRNA type	Anticodon	No. found	tRNA type	Anticodon	No. found	tRNA type	Anticodon	No. found	tRNA type	Anticodon	No. found
Phe	AAA	0	**Ser**	AGA	6	**His**	ATG	0	**Ile**	AAT	1
	GAA	2		GGA	0		GTG	3		GAT	0
Leu	TAA	1		TGA	2	**Gln**	TTG	3		TAT	1
	CAA	4		CGA	6		CTG	8	**Asp**	ATC	0
	AAG	2		ACT	0	**Asn**	ATT	0		GTC	9
	GAG	0		GCT	7		GTT	0	**Glu**	TTC	3
	TAG	2	**Thr**	AGT	3	**Lys**	TTT	6		CTC	2
	CAG	8		GGT	0		CTT	1	**Gly**	ACC	0
Val	AAC	0		TGT	4	**Cys**	ACA	0		GCC	3
	GAC	5		CGT	1		GCA	0		TCC	1
	TAC	3	**Tyr**	ATA	0	**Arg**	ACG	4		CCC	5
	CAC	2		GTA	3		GCG	0	**Met**	CAT	1
Pro	AGG	12	**Ala**	AGC	3		TCG	3	**Trp**	CCA	7
	GGG	0		GGC	0		CCG	2			
	TGG	5		TGC	3		TCT	3	**Total Anticodons found:**		
	CGG	5		CGC	5		CCT	1			161

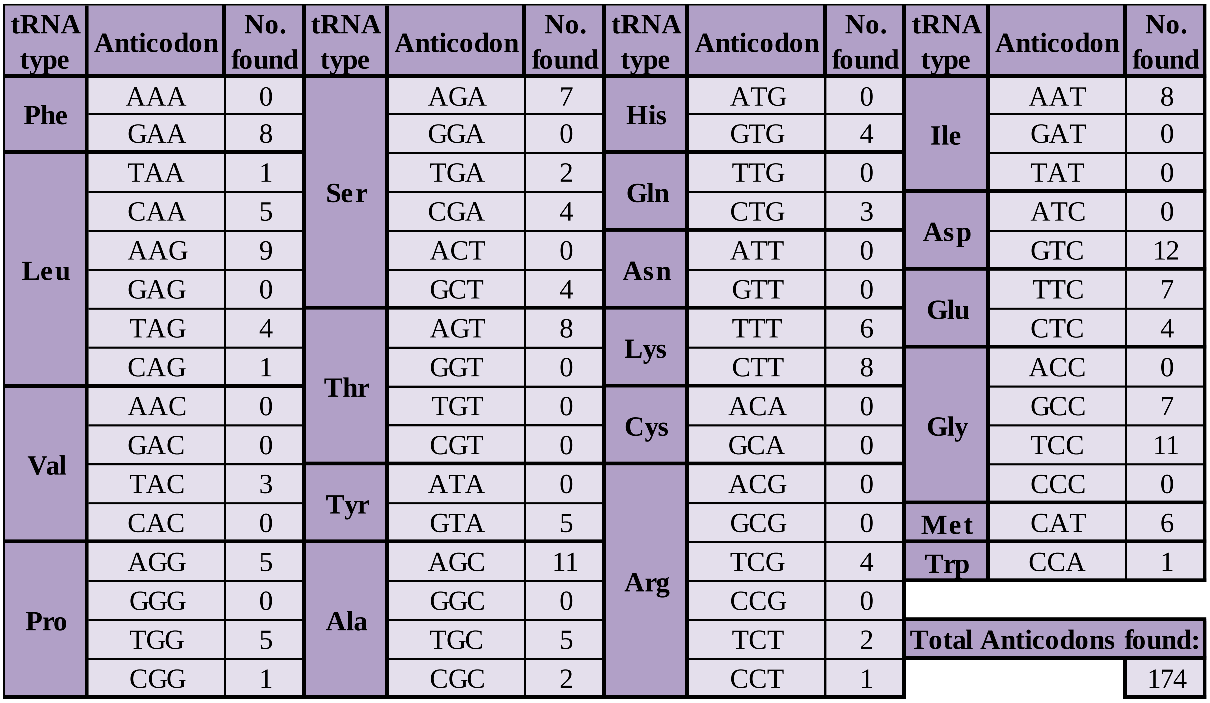
Table 199 | Number and type of tRNA genes found for the Stramenopile *A. astaci* (chimeric and pseudo genes removed)

tRNA type	Anticodon	No. found	tRNA type	Anticodon	No. found	tRNA type	Anticodon	No. found	tRNA type	Anticodon	No. found
Phe	AAA	0	**Ser**	AGA	7	**His**	ATG	0	**Ile**	AAT	2
	GAA	3		GGA	0		GTG	0		GAT	0
Leu	TAA	2		TGA	2	**Gln**	TTG	7		TAT	1
	CAA	4		CGA	14		CTG	7	**Asp**	ATC	0
	AAG	10		ACT	0	**Asn**	ATT	0		GTC	14
	GAG	0		GCT	11		GTT	2	**Glu**	TTC	2
	TAG	3	**Thr**	AGT	3	**Lys**	TTT	8		CTC	3
	CAG	10		GGT	0		CTT	2	**Gly**	ACC	0
Val	AAC	0		TGT	5	**Cys**	ACA	0		GCC	11
	GAC	8		CGT	1		GCA	0		TCC	0
	TAC	6	**Tyr**	ATA	0	**Arg**	ACG	9		CCC	6
	CAC	5		GTA	2		GCG	0	**Met**	CAT	1
Pro	AGG	2	**Ala**	AGC	5		TCG	4	**Trp**	CCA	9
	GGG	0		GGC	0		CCG	1			
	TGG	7		TGC	3		TCT	4	**Total Anticodons found:**		
	CGG	5		CGC	5		CCT	2			218

Table 200 | Number and type of tRNA genes found for the Stramenopile *A. invadans* (chimeric and pseudo genes removed)

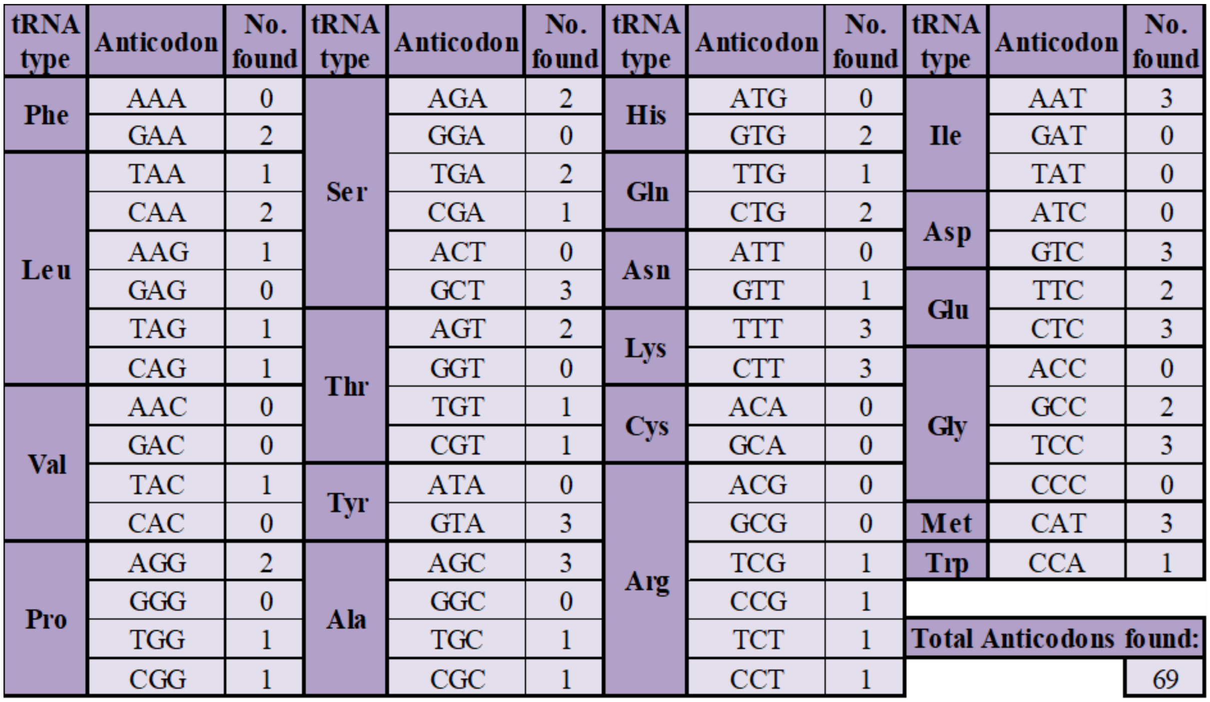
tRNA type	Anticodon	No. found	tRNA type	Anticodon	No. found	tRNA type	Anticodon	No. found	tRNA type	Anticodon	No. found
Phe	AAA	0	**Ser**	AGA	10	**His**	ATG	0	**Ile**	AAT	2
	GAA	5		GGA	0		GTG	2		GAT	0
Leu	TAA	1		TGA	2	**Gln**	TTG	3		TAT	2
	CAA	7		CGA	5		CTG	1	**Asp**	ATC	0
	AAG	5		ACT	0	**Asn**	ATT	0		GTC	6
	GAG	0		GCT	6		GTT	8	**Glu**	TTC	0
	TAG	2	**Thr**	AGT	7	**Lys**	TTT	3		CTC	1
	CAG	5		GGT	0		CTT	11	**Gly**	ACC	0
Val	AAC	1		TGT	3	**Cys**	ACA	0		GCC	1
	GAC	5		CGT	5		GCA	4		TCC	6
	TAC	2	**Tyr**	ATA	0	**Arg**	ACG	0		CCC	5
	CAC	2		GTA	4		GCG	0	**Met**	CAT	8
Pro	AGG	5	**Ala**	AGC	6		TCG	4	**Trp**	CCA	0
	GGG	0		GGC	0		CCG	4			
	TGG	3		TGC	7		TCT	2	**Total Anticodons found:**		
	CGG	3		CGC	0		CCT	1			175

Table 201 | Number and type of tRNA genes found for the Stramenopile *S. parasitica* (chimeric and pseudo genes removed)



tRNA type	Anticodon	No. found	tRNA type	Anticodon	No. found	tRNA type	Anticodon	No. found	tRNA type	Anticodon	No. found
Phe	AAA	0	**Ser**	AGA	7	**His**	ATG	0	**Ile**	AAT	8
	GAA	8		GGA	0		GTG	4		GAT	0
Leu	TAA	1		TGA	2	**Gln**	TTG	0		TAT	0
	CAA	5		CGA	4		CTG	3	**Asp**	ATC	0
	AAG	9		ACT	0	**Asn**	ATT	0		GTC	12
	GAG	0		GCT	4		GTT	0	**Glu**	TTC	7
	TAG	4	**Thr**	AGT	8	**Lys**	TTT	6		CTC	4
	CAG	1		GGT	0		CTT	8	**Gly**	ACC	0
Val	AAC	0		TGT	0	**Cys**	ACA	0		GCC	7
	GAC	0		CGT	0		GCA	0		TCC	11
	TAC	3	**Tyr**	ATA	0	**Arg**	ACG	0		CCC	0
	CAC	0		GTA	5		GCG	0	**Met**	CAT	6
Pro	AGG	5	**Ala**	AGC	11		TCG	4	**Trp**	CCA	1
	GGG	0		GGC	0		CCG	0			
	TGG	5		TGC	5		TCT	2	**Total Anticodons found:**		
	CGG	1		CGC	2		CCT	1			174

Table 202 | Number and type of tRNA genes found for the Stramenopile *P. multistriata* (chimeric and pseudo genes removed)



tRNA type	Anticodon	No. found	tRNA type	Anticodon	No. found	tRNA type	Anticodon	No. found	tRNA type	Anticodon	No. found
Phe	AAA	0	**Ser**	AGA	2	**His**	ATG	0	**Ile**	AAT	3
	GAA	2		GGA	0		GTG	2		GAT	0
Leu	TAA	1		TGA	2	**Gln**	TTG	1		TAT	0
	CAA	2		CGA	1		CTG	2	**Asp**	ATC	0
	AAG	1		ACT	0	**Asn**	ATT	0		GTC	3
	GAG	0		GCT	3		GTT	1	**Glu**	TTC	2
	TAG	1	**Thr**	AGT	2	**Lys**	TTT	3		CTC	3
	CAG	1		GGT	0		CTT	3	**Gly**	ACC	0
Val	AAC	0		TGT	1	**Cys**	ACA	0		GCC	2
	GAC	0		CGT	1		GCA	0		TCC	3
	TAC	1	**Tyr**	ATA	0	**Arg**	ACG	0		CCC	0
	CAC	0		GTA	3		GCG	0	**Met**	CAT	3
Pro	AGG	2	**Ala**	AGC	3		TCG	1	**Trp**	CCA	1
	GGG	0		GGC	0		CCG	1			
	TGG	1		TGC	1		TCT	1	**Total Anticodons found:**		
	CGG	1		CGC	1		CCT	1			69

Table 203 | Number and type of tRNA genes found for the Stramenopile *T. pseudonana* (chimeric and pseudo genes removed)

tRNA type	Anticodon	No. found	tRNA type	Anticodon	No. found	tRNA type	Anticodon	No. found	tRNA type	Anticodon	No. found
Phe	AAA	0	**Ser**	AGA	1	**His**	ATG	0	**Ile**	AAT	1
	GAA	2		GGA	0		GTG	1		GAT	0
Leu	TAA	0		TGA	2	**Gln**	TTG	0		TAT	0
	CAA	1		CGA	1		CTG	1	**Asp**	ATC	0
	AAG	1		ACT	0	**Asn**	ATT	0		GTC	1
	GAG	0		GCT	2		GTT	2	**Glu**	TTC	1
	TAG	1	**Thr**	AGT	0	**Lys**	TTT	1		CTC	1
	CAG	0		GGT	0		CTT	1	**Gly**	ACC	0
Val	AAC	0		TGT	1	**Cys**	ACA	0		GCC	1
	GAC	1		CGT	1		GCA	0		TCC	1
	TAC	1	**Tyr**	ATA	0	**Arg**	ACG	1		CCC	1
	CAC	1		GTA	2		GCG	0	**Met**	CAT	3
Pro	AGG	0	**Ala**	AGC	1		TCG	0	**Trp**	CCA	0
	GGG	0		GGC	0		CCG	0			
	TGG	0		TGC	4		TCT	0	***Total Anticodons found:***		
	CGG	1		CGC	1		CCT	0			42

Table 204 | Number and type of tRNA genes found for the Stramenopile *N. gaditana* (chimeric and pseudo genes removed)

tRNA type	Anticodon	No. found	tRNA type	Anticodon	No. found	tRNA type	Anticodon	No. found	tRNA type	Anticodon	No. found
Phe	AAA	0	**Ser**	AGA	1	**His**	ATG	0	**Ile**	AAT	1
	GAA	2		GGA	0		GTG	0		GAT	0
Leu	TAA	1		TGA	1	**Gln**	TTG	1		TAT	3
	CAA	1		CGA	2		CTG	0	**Asp**	ATC	0
	AAG	3		ACT	0	**Asn**	ATT	0		GTC	3
	GAG	0		GCT	1		GTT	2	**Glu**	TTC	1
	TAG	1	**Thr**	AGT	1	**Lys**	TTT	1		CTC	2
	CAG	0		GGT	0		CTT	3	**Gly**	ACC	0
Val	AAC	0		TGT	1	**Cys**	ACA	0		GCC	3
	GAC	0		CGT	2		GCA	0		TCC	1
	TAC	0	**Tyr**	ATA	0	**Arg**	ACG	4		CCC	1
	CAC	0		GTA	0		GCG	0	**Met**	CAT	0
Pro	AGG	2	**Ala**	AGC	0		TCG	2	**Trp**	CCA	1
	GGG	0		GGC	0		CCG	2			
	TGG	1		TGC	1		TCT	2	**Total Anticodons found:**		
	CGG	1		CGC	2		CCT	1			58

Table 205 | Number and type of tRNA genes found for the Stramenopile *A. anophagefferens* (chimeric and pseudo genes removed)

tRNA type	Anticodon	No. found	tRNA type	Anticodon	No. found	tRNA type	Anticodon	No. found	tRNA type	Anticodon	No. found
Phe	AAA	0	**Ser**	AGA	1	**His**	ATG	0	**Ile**	AAT	0
	GAA	2		GGA	0		GTG	1		GAT	0
Leu	TAA	1		TGA	1	**Gln**	TTG	2		TAT	1
	CAA	1		CGA	2		CTG	2	**Asp**	ATC	0
	AAG	4		ACT	1	**Asn**	ATT	0		GTC	3
	GAG	0		GCT	1		GTT	0	**Glu**	TTC	5
	TAG	1	**Thr**	AGT	2	**Lys**	TTT	2		CTC	2
	CAG	1		GGT	0		CTT	2	**Gly**	ACC	0
Val	AAC	1		TGT	2	**Cys**	ACA	0		GCC	2
	GAC	0		CGT	2		GCA	1		TCC	3
	TAC	2	**Tyr**	ATA	0	**Arg**	ACG	2		CCC	1
	CAC	1		GTA	2		GCG	0	**Met**	CAT	3
Pro	AGG	2	**Ala**	AGC	1		TCG	1	**Trp**	CCA	1
	GGG	0		GGC	0		CCG	1			
	TGG	2		TGC	3		TCT	1	**Total Anticodons found:**		
	CGG	0		CGC	0		CCT	1			73

Table 206 | Number and type of tRNA genes found for the Rhizarian *B. natans* (chimeric and pseudo genes removed)

tRNA type	Anticodon	No. found	tRNA type	Anticodon	No. found	tRNA type	Anticodon	No. found	tRNA type	Anticodon	No. found
Phe	AAA	0	**Ser**	AGA	3	**His**	ATG	0	**Ile**	AAT	6
	GAA	4		GGA	0		GTG	3		GAT	0
Leu	TAA	1		TGA	1	**Gln**	TTG	2		TAT	0
	CAA	2		CGA	4		CTG	0	**Asp**	ATC	0
	AAG	3		ACT	0	**Asn**	ATT	0		GTC	8
	GAG	0		GCT	2		GTT	4	**Glu**	TTC	4
	TAG	1	**Thr**	AGT	3	**Lys**	TTT	0		CTC	0
	CAG	5		GGT	0		CTT	7	**Gly**	ACC	0
Val	AAC	5		TGT	2	**Cys**	ACA	0		GCC	6
	GAC	0		CGT	3		GCA	3		TCC	2
	TAC	1	**Tyr**	ATA	0	**Arg**	ACG	5		CCC	2
	CAC	3		GTA	2		GCG	0	**Met**	CAT	5
Pro	AGG	3	**Ala**	AGC	5		TCG	0	**Trp**	CCA	2
	GGG	0		GGC	0		CCG	1			
	TGG	2		TGC	2		TCT	0	**Total Anticodons found:**		
	CGG	4		CGC	4		CCT	0			125

Table 207 | Number and type of tRNA genes found for the Rhizarian *P. brassicae* (chimeric and pseudo genes removed)