Codon Usage Bias and RNA Secondary Structures Analysis for Virus Resistant Genes in *Arabidopsis thaliana* and *Oryza sativa*

Muhammad Younas Khan Barozai¹, Syed Qasim Shah¹, Muhammad Din^{1*}, Raz Muhammad¹

¹Department of Botany, University of Balochistan, Sariab Road Quetta, Pakistan

*Corresponding author E-mail deen.2006@yahoo.com

Phone # +92-0333-7817319 Fax +92-081-9211-277

ABSTRACT

The current study is based upon the relationship of the codon usage bias (CUB) and RNA secondary structures (RSS) analysis between virus resistant (biotic stress) and housekeeping genes in *Arabidopsis thaliana* (*ath*) and *Oryza sativa* (*osa*). A total of seventeen genes (including thirteen virus resistant and four housekeeping genes) were subjected for the CUB and RSS analysis. The CUB analysis revealed that nine (out of thirteen) virus resistant genes showed similarity by more than 50% in the usage of codons for *Arabidopsis thaliana* and *Oryza sativa* and only one (out of four) ATC8 (housekeeping gene) showed similarity in the usage of codons by more than 50%. On the other hand Mfold algorithms for RNA secondary structure prediction revealed that six virus resistant genes showed significant difference in minimum free energy, two in number of loops and one in the number of stems. Whereas for the housekeeping genes, none showed significant difference in minimum free energy, number of loops & number of stems of RNA secondary structures. It can be concluded from the results that on the basis of RNA secondary structure virus resistant genes show mix behavior while housekeeping genes show uniform behavior. CUB analysis helps us to engineer the virus resistant crops through adjusting specific codon usage.

Keywords: Biotic stress; CUB; Housekeeping genes; RNA secondary structures;

Introduction

CUB refers to the nonrandom usage of synonymous codon to encode a specific amino acid (AA). Frequently used codons are termed as optimal or major codons, whereas less frequently used ones are termed non optimal or minor codons. Non optimal codons usually correspond to less abundant tRNAs than that of optimal codons [1,2,3]. Different organisms show particular preference to one of the several codons that encode the same amino acid. The arising of these preferences is much debated area of molecular evolution. It is generally acknowledged that codon preferences reflect a balance between mutational biases and natural selection for translational optimization [4]. In many unicellular organisms, invertebrates, and plants synonymous codon usage bias result from a coadaptation between codon usage and tRNAs abundance to optimize the efficiency of protein synthesis. However it remains unclear whether natural selection acts at the level of the speed or the accuracy of mRNAs translation. Codon usage can improve the fidelity of protein synthesis in multicellular species [5].

The consequences of biotic stress have been poorly understood because its application is difficult to control and its physiological consequences are highly variable. Many plant viruses are recognized on the basis of leaf symptoms that depend on localized changes to chloroplast structure and function. Virus infections may have greater effects on fitness and competitive ability in lower amount of Nitrogen (N), high light of the environments than that of shade,

high nutrient conditions. Some other ecological implications of these observations are also found to be responsible for the susceptibility of the plants to viral infection [6].

The first phase of resistance is induced by the recognition of pathogen-associated molecular patterns (PAMPs) by plant cell surface pattern recognition receptors, which initiates PAMP-triggered immunity that usually halts the infection of pathogens before invasion⁷⁻⁸. The next phase of plant resistance, resistance (R)-mediated resistance, or effector-triggered immunity, is induced by the direct or indirect recognition of pathogen effector proteins by plant R proteins, which are typically nucleotide binding site–Leu-rich repeat (NB-LRR) proteins [7,8]. Effector-triggered immunity usually induces a hypersensitive response (HR) with localized cell death and defense gene expression that suppresses the growth and spread of pathogens postentry [7,9].

The resistance to plant viruses can be divided into multiple stages [10]. The primary stage of virus resistance is the cellular-level resistance that occurs immediately after entry of the virus into plant cells which is called extreme resistance that inhibits viral accumulation in the initially invaded cells [10,11].

More than a dozen dominant genes responsible for resistance to plant viruses have been isolated, most of them are NB-LRR–type R genes [12]. RTM1 was isolated as the first lectin gene responsible for resistance to a Potyvirus [13]. But the importance of lectins in plant immunity to viruses has been debated for more than a decade. The lectin gene JAX1, targets

potexviruses, which are distantly related to potyviruses. JAX1 also exhibits a level of resistance different to that of RTM1.

Interestingly, out of the 14 natural recessive resistance genes against plant viruses that have been cloned from diverse plant species thus far, 12 encode the eukaryotic translation initiation factor 4E (eIF4E) or its isoform eIF(iso)4E [14].

In this study 13 virus resistant genes i.e., Resistance to yellow strain of cucumber mosaic virus1 (RCY1) [15], Enhanced disease susceptibility1 (EDS1) [16], Enhanced disease susceptibility5 (EDS5) [17], Restricted TEV movement (RTM1) [18], loss of susceptibility to Potyvirus 1 (LSP1) [19], Suppressor of gene silencing2 (SGS2) [20], Suppressor of gene silencing3 (SGS3) [20], eukaryotic initiation factor 4E1 (At-eIF4E1) [21], eukaryotic initiation factor (iso)4E (At-eIF(iso)4E) [22], Cucumovirus multiplication1 (cum1) [23], Cucumovirus multiplication2 (cum2) [23], Restricted TEV movement2 (RTM2) [24], hypersensitive response (HRT) [25] and four housekeeping genes Actin8 (ATC8) [26], Polyubiquitin 10 (UBQ10) [26], tubulin beta 2 (TUB2) [26] Elongation factor-1 alpha (EF-1 α) [26] were selected from *Arabidopsis thaliana* and *Oryza sativa* and subjected to CUB analysis, using bioinformatics tools. The aim of the study is to find co-relation between CUB and virus stress along with the housekeeping genes in *Arabidopsis thaliana* and *Oryza sativa*.

The secondary structure of a nucleic acid molecule refers to the base pairing interactions within a single molecule or set of interacting molecules, & can be represented as a list of bases which are paired in a nucleic acid molecule [27]. The secondary structures of biological DNA's and RNA's tend to be different: biological DNA mostly exists as fully base paired double helices, while biological RNA is single stranded & often forms complicated base-pairing interactions due to its increased ability to form hydrogen bonds stemming from the extra hydroxyl group in the ribose sugar.

Material and Methods

Identification of Virus Resistant Genes

Thirteen virus resistant genes (RCY1 [15], EDS1 [16], EDS5 [17], RTM1 [18], LSP1 [19], SGS3 [20], SGS2 [20], At-eIF4E1 [21], At-eIF (iso) 4E [22], cum1 [23], cum2 [23], RTM2 [24], HRT [25]) and four housekeeping genes (ACT8 [26], TUB2 [26], UBQ10 [26] and EF-1 α [26]) were identified in *Arabidopsis thaliana* and *Oryza sativa* through literature survey.

Sequence Retrieval

Thirteen virus resistant genes (RCY1, EDS1, EDS5, RTM1, LSP1, SGS2, SGS, At-eIF4E1, At-eIF (iso) 4E, cum1, cum2, RTM2, HRT) and four

housekeeping genes (ACT8, TUB2, UBQ10 and EF-1 α) in *Arabidopsis thaliana* and *Oryza sativa* were retrieved through nucleotide database and gen bank of NCBI by using the accession number of the genes. The fasta sequences of genes were saved and used for further analysis of RNA secondary structures through Mfold algorithm.

Open Reading Fram (ORF) Analysis

Open reading frame (ORF) is a graphical analysis tool which is used to find all the open reading frames of a nucleotide sequence. The ORF feature for all viral resistant genes and housekeeping genes were found through ORF finder. The ORF finder plays an important role in the determination of protein sequence of the nucleotides. The fasta sequences of virus resistant and housekeeping genes of *Arabidopsis thaliana* and *Oryza sativa* were subjected to the ORF finder of NCBI and the longest frames were selected for further analysis.

Secondary structure analysis

Secondary structure of RNAs were predicted from Mfold [28] and then analyzed i.e. the number of loops, budges, junction & helices, number of stems, loops, & minimum free energy of each RNA secondary structure was recorded.

Codon Usage Bias Analysis

For the analysis of codon usage the fasta sequence of the genes were subjected to codon usage analysis program i.e., www.geneinfinity.org. The codon usage analysis program was used for analysis of amino acids. The codon having largest value were selected and compared them for each amino acid of *Ath* and *Osa*. It is the key step of research and the results obtained in this step were saved for further analysis.

The relative synonymous codon usage (RSCU) values for each of the codon in the sequence were determined by using the following formula.

$$RSCU = (\text{no. of codon used} / \text{total no. of codon}) \times \text{no. of amino acids.}$$

Results

Virus Resistant Genes in term of open reading frame (ORF)

Thirteen genes (RCY1, EDS1, EDS5, RTM1, LSP1, SGS2, SGS, At-eIF4E1, At-eIF (iso) 4E, cum1, cum2, RTM2, HRT) were identified as virus resistant genes and four as housekeeping genes (ACT8, TUB2, UBQ10 and EF-1 α) in *Arabidopsis thaliana* and *Oryza sativa*. The ORF analysis (Table 1) reveals that For RCY1 gene *Arabidopsis thaliana* (*Ath*) prefers frame +1 and *Oryza sativa* (*Osa*) prefers +3 frame and the frame length in base pairs (bp) are 1086 and 1284 respectively. While frame length of amino acid (A.A) are 361 and 427 respectively similarly for EDS1 +1 and +2 and frame length of bp and A.A are 1548, 1866 and 515, 621. EDS5 prefer the same frame i.e., +2 in both plants and frame length of bp and A.A are

1632,414. RTM1 prefer +2 and +1 fram and fram length of bp and A.A are 525,483 and 174,160. LSP1 prefer the same fram of +2 while 504,621 and 167,206 are fram length of bp and A.A in *Ath* and *Osa* respectively. SGS2 prefer +1 fram in cases and 2805, 3657 and 934, 1218 are fram length of bp and A.A respectively. In SGS3 *Ath* prefer +2 and that of *Osa* prefer +3 and 1878,1830 are fram length of bp while 625,609 are fram length of A.A.eIF4E prefer +3 fram and fram length of bp, A.A are 708,684 and 235,227 respectively. In elf (iso)4E the two plants prefer the fram of +1 and +2 and their fram length of bp and A.A are 366,621 and 121,206 respectively.CUM1 gene in *Ath* and *Osa* prefer +1 and -2 while fram length of bp and A.A are 708,360 and 235,119. In case of CUM2 gene *Ath* and *Osa* prefer +2and +3 fram and 5178, 2445 are fram length of bp while 1725,814 are fram length of A.A respectively. In RTM2 gene plants prefer +3 fram, the fram length in bp of *Ath* is 1101 and in case of *Osa* it is 1017, and that of A.A are 366 and 338 respectively. The HRT gene *Ath* prefer +1 and *Osa* prefer +2 fram, their length of bp are 1683, 2700 and that of A.A are 560,899 respectively.

The gene length of *Ath* genes are 1582, 2036, 2031, 634, 1293, 4013, 2145, 942, 1474, 898, 5456, 1209, 5996 and that of *Osa* are 1841, 2140, 7337, 3606, 1124, 2027, 958, 3932, 2177, 750, 993, 3227, 2942, 1432 and 3454 base pairs respectively.

The ORF analysis were also done for four housekeeping genes (Table 1) in which for ATC8 gene both *Ath* and *Osa* prefer +1 fram, 1134 fram length of bp and 377 fram length of A.A respectively. For UBQ 10 gene *Ath* and *Osa* prefer +3 and +1 their fram length of bp is 1146 and 1209 and that of A.A are 381 and 402 respectively. *Ath* prefer +2 fram for TUB2 gene and *Osa* prefer +1 and fram length of bp are 1353, 1344 and that of A.A are 450,447 respectively. For EF-1 α gene *Ath* and *Osa* prefer +3 and +2 fram and 1350, 1344 are fram length of bp and 449,447 are of A.A respectively.

The gene length of *Ath* housekeeping gene are 1707, 1270, 1726, 1590 and that of *Osa* gene are 1608, 1609, 1799 and 1755 base pairs respectively.

Codon Usage Bias (CUB) Analysis

The CUB analysis for thirteen virus resistant (RCY1, EDS1, EDS5, RTM1, LSP1, SGS2, SGS, At-eIF4E1, At-eIF (iso) 4E, cum1, cum2, RTM2, HRT) and four housekeeping genes (ACT8, TUB2, UBQ10 and EF-1 α) in *Arabidopsis thaliana* and *Oryza sativa* showed interesting result. The (Table 2(a, b, c) and 3) summarize the mostly used codon usage bias (CUB) calculated in term of Relative synonymous codon usage (RSCU) for the thirteen selected virus resistant and four housekeeping genes in *Arabidopsis*

thaliana and *Oryza sativa* for the 20 amino acids and stop codon.

The nine (9) virus resistant genes show similarity more than 50% i.e., RCY1 12 codons (57.14%), EDS1 13 codons (61.90%), EDS5 18 codons (85.71%). RTM1 13 codons (61.90%), SGS3 11 codons (52.38%), eIF4E 14 codon (66.66%), elf (iso) 4E 13 codons (61.90%), CUM1 13 codons (61.90%) and that of HRT show similarity in the usage of 20 codons (95.23%) while the other 4 genes showed similarity by less than 50% i.e., LSP1 7 codons (33.33%), SGS2 6 codons (28.57%), CUM2 9 codons (42.85%) and RTM2 10 codons (47.61%). In housekeeping genes only ACT8 show similarity in the usage of codons by more than 50% i.e., 15 codons (71.42%) while remaining 3 genes show similarities by less than 50% i.e., UBQ10 7 codons (33.33%), TUB28 codons (38.09%), and EF-1 α 10 codons (47.61%) in *Arabidopsis thaliana* and *Oryza sativa*. The 69% virus resistant and 25% housekeeping genes showing >50% similarity, suggests the straight correlation between the CUB: virus stress and CUB: housekeeping, irrespective of the plant species.

Secondary structure Analysis

The RNA secondary structures of virus resistant and housekeeping genes were analyzed (Table 4.) in terms of Number of loops, stem region with Minimum free energy (MFE). Six of the virus resistant genes (RCY1, EDS1, LSP1 CUM1, RTM2 and HRT) show significant difference in minimum free energy, two (CUM1 and RTM2) Show significant difference in number of loops and only one (RTM2) show significant difference for the number of stems in secondary structures. Whereas for the house keeping genes, no one showed significant difference in minimum free energy, number of loops & number of stems. It can conclude from the results that mix behavior on the basis of RNA secondary structures were observed for viral resistant genes and uniform for housekeeping genes in *Arabidopsis thaliana* & *Oryza sativa*.

Discussions

The study is based upon codon usage bias in *Arabidopsis thaliana* and *Oryza sativa* through bioinformatics tools. Bioinformatics is the part of molecular biology that involves working with biological data, typically using computers, with the goal of enabling and accelerating biological research. Bioinformatics spans a wide range of activities: data capture, automated recording of experimental results; data storage and access, using a multitude of databases and query tools; data analysis; and visualization of raw data and analytical results. Bioinformatics is also applied in the creation and

Table # 1. ORF Features for Virus Resistant (VR) Genes and Housekeeping (HK) Genes in *Ath* and *Osa* .

S. #	Gene Nature	Gene Name	Gene length		Frame preferred		Frame length (bp)		Frame length (A.A)	
			<i>Ath</i>	<i>Osa</i>	<i>Ath</i>	<i>Osa</i>	<i>Ath</i>	<i>Osa</i>	<i>Ath</i>	<i>Osa</i>
1	VR	RCY1	1582 bp	1841 bp	+ 1	+3	1086	1284	361	427
2	VR	EDS1	2036 bp	2140 bp	+1	+2	1548	1866	515	621
3	VR	EDS5	2031 bp	7337 bp	+2	+2	1632	414	543	137
4	VR	RTM1	634 bp	3606 bp	+2	+1	525	483	174	160
5	VR	LSP1	1293 bp	958 bp	+2	+2	504	621	167	206
6	VR	SGS2	4013 bp	3932 bp	+1	+1	2805	3657	934	1218
7	VR	SGS3	2145 bp	2177 bp	+2	+3	1878	1830	625	609
8	VR	elF4E	942 bp	750 bp	+3	+3	708	684	235	227
9	VR	elF(iso)4E	1474 bp	993 bp	+1	+2	366	621	121	206
10	VR	CUM1	898 bp	3227 bp	+1	-2	708	360	235	119
11	VR	CUM2	5456 bp	2942 bp	+2	+3	5178	2445	1725	814
12	VR	RTM2	1209 bp	1432 bp	+3	+3	1101	1017	366	338
13	VR	HRT	5996 bp	3454 bp	+1	+2	1683	2700	560	899
14	HK	ACT8	1707 bp	1608 bp	+1	+1	1134	1134	377	377
15	HK	UBQ10	1270 bp	1609 bp	+3	+1	1146	1209	381	402
16	HK	TUB2	1726 bp	1799 bp	+2	+1	1353	1344	450	447
17	HK	EF1- α	1590 bp	1755 bp	+3	+2	1350	1344	449	447

Table #.2 (a) RSCU (relative synonymous codon usage) Values for Each Amino Acid of Virus Resistant Genes in *Ath* and *Osa*.

Amino acids	RCY1		EDS 1		EDS 5		RTM1	
	<i>A.thaliana</i>	<i>O.sativa</i>	<i>A.thaliana</i>	<i>O.sativa</i>	<i>A.thaliana</i>	<i>O.sativa</i>	<i>A.thaliana</i>	<i>O.sativa</i>
Ala	GCC(1.47)	GCG(1.30)	GCT(1.66)	GCA,GCC(1.11)	GCT,GCC(1.52)	GCT(1.54)	GCG,GCA,GCC,GCT(0.00)	GCC(1.57)
Cys	TGT(1.07)	TGC(1.17)	TGT(1.26)	TGC(1.08)	TGC(1.22)	TGC(1.09)	TGT,TGC(1.00)	TGT(1.11)
Asp	GAT(1.27)	GAT,GAC(1.00)	GAT(1.52)	GAC(1.05)	GAT(1.78)	GAT(1.50)	GAT(1.2)	GAT(1.06)
Glu	GAA(1.65)	GAG(1.33)	GAA(1.34)	GAG,GAA(1.00)	GAA(1.14)	GAA(1.59)	GAA(1.25)	GAA(1..21)
Phe	TTC(1.08)	TTC(1.06)	TTT(1.17)	TTC(1.28)	TTT(1.13)	TTT(1.35)	TTT(1.33)	TTT(1.38)
Gly	GGA(3.71)	GGG(1.30)	GG(1.88)	GGA(1.49)	GGT(1.19)	GGA(1.48)	GGG,GGA(2.00)	GGC(1.18)
His	CAT(1.33)	CAT,CAC(1.00)	CAT,CAC(1.00)	CAT(0.92)	CAT(1.37)	CAT(1.38)	CAT(1.33)	CAT(1.22)
Ile	ATT(1.44)	ATA(1.42)	ATT(1.50)	ATT(1.28)	ATT(1.54)	ATT(1.34)	ATT(1.80)	ATT(1.25)
Lys	AAG(1.18)	AAG(1.17)	AAA(1.33)	AAG(1.50)	AAG(1.14)	AAA(1.21)	AAA(1.81)	AAG(1.08)
Leu	TTG,CTA(1.26)	TTG(1.71)	CTT(1.88)	CTT(1.60)	CTC(1.92)	CTT(1.45)	TTA,CTG(1.84)	TTA,CTT(1.38)
Met	ATG(1.00)	ATG(1.00)	ATG(1.00)	ATG(1.00)	ATG(1.00)	ATG(1.00)	ATG(0.00)	ATG(1.00)
Asn	AAT(1.14)	AAC(1.05)	AAT(1.21)	AAT,AAC(1.00)	AAC(1.08)	AAC(1.02)	AAT(1.33)	AAT(1.45)
Pro	CCT,CCC(1.11)	CCG,CCA,CCC(1.06)	CCA(1.68)	CCA,CCC(1.11)	CCA,CCT(1.33)	CCT(1.64)	CCA(2.40)	CCC(1.16)
Gln	CAA(1.07)	CAA(1.09)	CAA(1.41)	CAA(1.25)	CAA(1.57)	CAA(1.13)	CAA(1.84)	CAA(1.20)
Arg	AGA(2.94)	AGA(2.03)	AGA(3.03)	CGC(2.05)	AGA(1.86)	AGA(1.66)	CGG(1.71)	CGC(2.44)
Ser	TCA(1.47)	TCC(1.63)	AGC,TCA(1.90)	AGC(0.95)	TCT(1.79)	TCT(1.90)	TCA,TCT(1.71)	TCC(1.48)
Thr	ACT(1.45)	ACT(1.33)	ACA(1.81)	ACT(1.77)	ACT(1.6)	ACT(3.84)	ACT(2.28)	ACA(1.88)
Val	GTT(1.62)	GTG(1.92)	GTT(2.12)	GTC(1.50)	GTT(2.00)	GTT(1.68)	GTA(2.00)	GTA(1.22)
Trp	TGG(1.00)	TGG(1.00)	TGG(1.00)	TGG(1.00)	TGG(1.00)	TGG(1.00)	TGG(1.00)	TGG(1.00)
Tyr	TAC(1.12)	TAT(1.09)	TAT(1.22)	TAT,TAC(1.00)	TAT(1.18)	TAT(1.33)	TAC(1.11)	TAT(1.21)
End	TGA(1.37)	TGA(1.75)	TGA(1.76)	TGA(1.80)	TAA(1.69)	TAA(1.26)	TGA(1.80)	TGA(1.33)

Table # 2. (b) RSCU (relative synonymous codon usage) Values for Each Amino Acid of Viral Resistant Genes in *Ath* and *Osa*

Amino acids	LSP1		SGS2		SGS3		EIF4E	
	<i>A.thaliana</i>	<i>O.sativa</i>	<i>A.thaliana</i>	<i>O.sativa</i>	<i>A.thaliana</i>	<i>O.sativa</i>	<i>A.thaliana</i>	<i>O.sativa</i>
Ala	GCT(1.69)	GCC(1.11)	GCT(1.78)	GCG(1.17)	GCA(2.31)	GCA(13.33)	GCA(2.85)	GCA(1.33)
Cys	TGT(1.50)	TGT(1.14)	TGT(1.15)	TGC(1.33)	TGT(1.17)	TGC(1.42)	TGT(1.33)	TGC(1.33)
Asp	GAT(1.33)	GAC(1.06)	GAT(1.42)	GAC(1.05)	GAT(1.58)	GAC(1.33)	GAT,GAC(1.00)	GAT(2.00)
Glu	GAG,GAA(1.00)	GAG(1.07)	GAA(1.19)	GAG(1.41)	GAA(1.19)	GAA(1.17)	GAG(1.40)	GAA(1.11)
Phe	TTT(1.36)	TTT(1.71)	TTT(1.16)	TTC(1.34)	TTT,TTC(1.00)	TTC(1.66)	TTT(1.38)	TTT(1.20)
Gly	GGA(1.84)	GGC(1.62)	GGG(1.45)	GGG(1.55)	GGA(2.36)	GGA(1.43)	GGA(2.66)	GGG,GGA(1.41)
His	CAC(1.50)	CAC(1.11)	CAT(1.40)	CAT,CAC(1.00)	CAT(1.52)	CAC(1.13)	CAT(2.00)	CAT(1.33)
Ile	ATA(1.61)	ATT(2.14)	ATT(1.32)	ATA(1.11)	ATC(1.50)	ATA(1.10)	ATC(1.28)	ATA(1.50)
Lys	AAG(1.30)	AAA(1.40)	AAA(1.01)	AAG(1.39)	AAA(1.28)	AAG(1.06)	AAG(1.33)	AAG,AAA(1.00)
Leu	TTG(1.86)	CTT(2.00)	TTG(1.57)	CTC(1.63)	CTT,TTA(1.33)	CTG(1.77)	TTG(1.46)	TTG(1.89)
Met	ATG(1.00)	ATG(1.00)	ATG(1.00)	ATG(1.00)	ATG(1.00)	ATG(1.00)	ATG(1.00)	ATG(1.00)
Asn	AAT(1.25)	AAC(1.33)	AAT(1.23)	AAC(1.05)	AAT(1.28)	AAC(1.10)	AAT(1.12)	AAT(1.6)
Pro	CCA(2.00)	CCT(1.47)	CCT(1.64)	CCG(1.28)	CCA(1.90)	CCG,CCC(1.21)	CCG(2.50)	CCG(1.76)
Gln	CAA(1.50)	CAA(1.40)	CAG(1.11)	CAG(1.13)	CAA(1.60)	CAA(1.07)	CAA(1.40)	CAA(1.33)
Arg	AGA(2.80)	CGG(2.05)	AGA(2.06)	AGG(2.10)	AGA(2.02)	AGA(2.43)	AGA(2.48)	CGC(1.94)
Ser	AGT(1.50)	AGC(1.44)	TCT(1.67)	TCC(1.29)	TCA(1.54)	AGC(1.71)	AGT(1.45)	AGC(1.40)
Thr	ACT(2.4)	ACG,ACC(1.33)	ACA(1.33)	ACC(1.10)	ACA(1.75)	ACA(1.82)	ACA(2.08)	ACG(1.38)
Val	GTC(1.60)	GTG(1.60)	GTT(1.31)	GTG(1.72)	GTT(2.00)	GTG(1.78)	GTT(1.84)	GTG(2.00)
Trp	TGG(1.00)	TGG(1.00)	TGG(1.00))	TGG(1.00)	TGG(1.00)	TGG(1.00)	TGG(1.00)	TGG(1.00)
Tyr	TAT(1.33)	TAC(1.50)	TAT(1.13)	TAC(1.35)	TAC(1.20)	TAT,TAC(1.00)	TAT(2.00)	TAT(2.00)
End	TAA(2.4)	TGA(1.95)	TGA(1.66)	TGA(2.25)	TGA(1.61)	TGA(1.85)	TGA(1.75)	TGA,TAG,TAA(1.00)

Table # 2. (c) RSCU (relative synonymous codon usage) Values for Each Amino Acid of Viral Resistant Genes in *Ath* and *Osa*.

Amino acids	eF(iso)4E		CUM1		CUM2		RTM2		HRT	
	<i>A.thaliana</i>	<i>O.sativa</i>	<i>A.thaliana</i>	<i>O.sativa</i>	<i>A.thaliana</i>	<i>O.sativa</i>	<i>A.thaliana</i>	<i>O.sativa</i>	<i>A.thaliana</i>	<i>O.sativa</i>
Ala	GCT(2.46)	GCT(1.41)	GCT(2.35)	GCC(1.63)	GCA(1.61)	GCA(1.69)	GCT(2.66)	GCG,GCC(1.33)	GCT(1.64)	GCT(1.68)
Cys	TGT(1.52)	TGT(1.06)	TGT(1.42)	TGC(1.02)	TGC(1.15)	TGT(1.02)	TGT(1.14)	TGC(1.37)	TGT(1.23)	TGT(1.16)
Asp	GAT(1.60)	GAT,CAC(1.00)	GAT(1.33)	GAT(1.36)	GAT(1.22)	GAT(1.12)	GAT(1.20)	GAC(1.25)	GAT(1.57)	GAT(1.45)
Glu	GAA(1.08)	GAG(1.20)	GAA(1.18)	GAA(1.33)	GAA(1.37)	GAG(1.12)	GAA(1.17)	GAG(1.26)	GAA(1.14)	GAA(1.12)
Phe	TTT(1.47)	TTT(1.71)	TTC(1.28)	TTT(1.50)	TTC(1.33)	TTT(1.39)	TTC(1.75)	TTC(1.03)	TTT(1.24)	TTT,TTC(1.00)
Gly	GGA,GGT(1.20)	GGC(1.62)	GGA(3.33)	GGA(1.15)	GGA(1.49)	GGG,GGC(1.13)	GGA(3.05)	GGA(1.55)	GGA(2.02)	GGA(2.14)
His	CAT(1.166)	CAT,CAC(1.00)	CAT(1.63)	CAT(1.39)	CAT(1.21)	CAC(1.16)	CAT(1.50)	CAT(1.07)	CAT(1.65)	CAT(1.52)
Ile	ATA(1.59)	ATT(2.14)	ATT(1.83)	ATT(1.14)	ATC(1.35)	ATT(1.11)	ATA(1.75)	ATC(1.17)	ATT(1.57)	ATT(1.5)
Lys	AAG(1.12)	AAA(1.40)	AAG(1.16)	AAA(1.63)	AAA(1.38)	AAG(1.13)	AAG(1.46)	AAG(1.02)	AAA(1.20)	AAG(1.06)
Leu	CTT(1.64)	CTT(2.06)	TTG(2.10)	CTC(1.53)	CTC(2.43)	TTG(1.90)	CTT(1.55)	TTG(1.92)	CTT(2.10)	CTT(1.30)
Met	ATG(1.00)	ATG(1.00)	ATG(1.00)	ATG(1.00)	ATG(1.00)	ATG(1.00)	ATG(1.00)	ATG(1.00)	ATG(1.00)	ATG(1.00)
Asn	AAT(1.21)	AAT(1.33)	AAC(1.05)	AAT(1.62)	AAT(1.06)	AAT(1.33)	AAT(1.15)	AAT(1.11)	AAT(1.08)	AAT(1.22)
Pro	CCT(2.00)	CCT(1.52)	CCT(2.54)	CCT(1.31)	CCA(1.46)	CCA(1.92)	CCA(2.00)	CCA(1.45)	CCT(1.71)	CCT(1.53)
Gln	CAG,CAA(1.00)	CAA(1.42)	CAG,CAA(1.00)	CAA(1.23)	CAA(1.47)	CAA(1.19)	CAA(1.06)	CAG,CAA(1.00)	CAA(1.48)	CAA(1.22)
Arg	CGA(1.84)	CGG(2.32)	AGA(2.57)	AGA,CGC(1.35)	AGA(2.57)	AGG(2.56)	AGA(2.70)	AGG(2.48)	AGA(3.32)	AGA(1.91)
Ser	TCT(2.10)	AGC,TCT(1.33)	TCA(1.50)	TCC(1.38)	TCC(1.49)	AGC(1.52)	AGC(1.90)	AGT(1.29)	AGT,TCT(1.25)	TCT(1.66)
Thr	ACT(2.00)	ACG,ACC(1.33)	ACT(1.86)	ACT(1.35)	ACC(1.70)	ACA(1.88)	ACG,ACA(1.20)	ACA(1.33)	ACA(1.89)	ACA,ACT(1.22)
Val	GTT(2.43)	GTG(1.60)	GTT(2.00)	GTT(1.23)	GTC(1.80)	GTG(1.58)	GTT(1.77)	GTG(1.88)	GTT(1.92)	GTT(1.58)
Trp	TGG(1.00)	TGG(1.00)	TGG(1.00)	TGG(1.00)	TGG(1.00)	TGG(1.00)	TGG(1.00)	TGG(1.00)	TGG(1.00)	TGG(1.00)
Tyr	TAT(1.36)	TAC(1.50)	TAC(1.33)	TAT(1.22)	TAT(0.96)	TAT(1.47)	TAT(1.33)	TAC(1.25)	TAT(1.39)	TAT(1.33)
End	TGA(1.83)	TGA(1.95)	TGA,TAG,TAA(1.00)	TGA(1.23)	TGA(1.77)	TGA(1.63)	TAA(1.22)	TGA(1.58)	TGA(1.83)	TGA(1.50)

Table # 3. RSCU (relative synonymous codon usage) Values for Each Amino Acid of Housekeeping Genes of *Ath* and *Osa*

Amino acids	ATC8		UBQ10		TUB2		EF1-alpha	
	<i>A.thaliana</i>	<i>O.sativa</i>	<i>A.thaliana</i>	<i>O.sativa</i>	<i>A.thaliana</i>	<i>O.sativa</i>	<i>A.thaliana</i>	<i>O.sativa</i>
Ala	GCT(1.88)	GCC(1.36)	GCT(2.00)	GCC(2.09)	GCA(1.83)	GCC(1.60)	GCG,GCA,GCC(1.14)	GCT(1.29)
Cys	TGT(1.40)	TGC(1.23)	TGC(1.50)	TGT(1.50)	TGT(1.07)	TGC(1.18)	TGC(1.33)	TGC(1.22)
Asp	GAT(1.15)	GAT(1.44)	GAT(1.33)	GAC(1.45)	GAT(1.52)	GAT(1.15)	GAT(1.09)	GAT(1.28)
Glu	GAG(1.31)	GAG(1.35)	GAA(1.62)	GAG(1.87)	GAA(1.20)	GAG(1.88)	GAG(1.42)	GAA(1.51)
Phe	TTC(1.09)	TTC(1.16)	TTC(1.52)	TTT(1.14)	TTT(1.04)	TTC(1.31)	TTC(1.20)	TTC(1.14)
Gly	GGT(2.70)	GGT(2.00)	GGC(2.00)	GGC(2.56)	GGA(1.94)	GGT,GGC(1.30)	GGT(1.86)	GGA(1.30)
His	CAT,CAC(1.00)	CAT(1.38)	CAC(2.00)	CAC(2.00)	CAT(1.07)	CAC(1.16)	CAT(1.25)	CAT(1.05)
Ile	ATT(1.42)	ATT(1.50)	ATA(1.66)	ATC(2.57)	ATC(1.28)	ATC(1.50)	ATT(1.38)	ATC(2.10)
Lys	AAG(1.03)	AAG(1.23)	AAG(1.17)	AAG(1.90)	AAA(1.64)	AAG(1.47)	AAG,AAA(1.00)	AAA(1.71)
Leu	CTT(2.37)	CTT(1.71)	TTG(1.46)	CTG(1.35)	CTT(1.67)	TTG(1.68)	TTG(2.21)	CTC(1.60)
Met	ATG(1.00)	ATG(1.00)	ATG(1.00)	ATG(1.00)	ATG(1.00)	ATG(1.00)	ATG(1.00)	ATG(1.00)
Asn	AAC(1.14)	AAC(1.33)	AAC(1.33)	AAC(1.50)	AAC(1.20)	AAC(1.33)	AAT(1.33)	AAC(1.33)
Pro	CCT(1.65)	CCT(1.57)	CCA(1.87)	CCC(1.50)	CCA(1.55)	CCT(1.53)	CCA(2.00)	CCA(1.41)
Gln	CAA(1.20)	CAA(1.09)	CAA(2.00)	CAG(1.90)	CAA(1.54)	CAG(1.52)	CAG(1.20)	CAA(1.86)
Arg	AGA(2.18)	AGG,AGA(1.35)	AGA(2.91)	CGT(2.89)	AGA(1.68)	CGT(2.06)	AGA(3.14)	CGT(2.04)
Ser	TCA(1.41)	AGC(1.53)	TCA,TCC(1.30)	TCT(1.93)	TCT(2.35)	TCT(1.95)	TCA(2.14)	TCT(2.05)
Thr	ACA(1.25)	ACT(1.62)	ACA(2.05)	ACC(3.06)	ACC(1.50)	ACT(1.41)	ACA(1.63)	ACG,ACT,ACC(1.33)
Val	GTT(2.10)	GTG(1.40)	GTA(1.48)	GTC(1.75)	GTT(1.67)	GTT(1.14)	GTG(1.48)	GTT(1.62)
Trp	TGG(1.00)	TGG(1.00)	TGG(1.00)	TGG(1.00)	TGG(1.00)	TGG(1.00)	TGG(1.00)	TGG(1.00)
Tyr	TAC(1.11)	TAT(1.61)	TAT(2.00)	TAC(1.55)	TAT(1.06)	TAC(1.47)	TAT(2.00)	TAT(1.71)
End	TGA,TAG(1.20)	TGA(1.67)	TGA(2.35)	TGA(1.80)	TGA(2.02)	TGA,TAG,TAA(1.00)	TGA(2.64)	TGA(2.41)

Table # 4. Comparison table of *Arabidopsis thaliana* & *Oryza sativa*.

The table describes in term of Plant Name, Accession number, Minimum Free Energy (MFE), Number of loops and Number of stems

1. RCY1 (resistance to CMV(y)1)				
Plant Name	Accession #	MFE(kcal/mol)	Number of Loops	Number of Stems
<i>Arabidopsis thaliana</i>	NM_128200	-185.77	27	26
<i>Oryza sativa</i>	NM_001049618	-266.23	25	23
2. EDS1 (Enhanced disease susceptibility1) gene				
Plant Name	Accession #	MFE(kcal/mol)	Number of Loops	Number of Stems
<i>Arabidopsis thaliana</i>	NM_114678	-184.91	30	31
<i>Oryza sativa</i>	NM_001069621	-290.01	30	27
3. EDS5 (Enhanced disease susceptibility5 gene)				
Plant Name	Accession #	MFE(kcal/mol)	Number of Loops	Number of Stems
<i>Arabidopsis thaliana</i>	AF416569	-224.66	30	28
<i>Oryza sativa</i>	NM_001052259	-272.35	32	28
4. RTM1 (Restricted TEV movement 1)				
Plant Name	Accession #	MFE(kcal/mol)	Number of Loops	Number of Stems
<i>Arabidopsis thaliana</i>	AF191302	-174.10	30	27
<i>Oryza sativa</i>	NM_001056884	-220.16	29	28
5. LSP1 (loss of susceptibility to Potyvirus 1)				
Plant Name	Accession #	MFE(kcal/mol)	Number of Loops	Number of Stems
<i>Arabidopsis thaliana</i>	NM_203123	-191.49	28	26
<i>Oryza sativa</i>	NM_001071345	-261.87	34	27
6. SGS2 (Suppressor of gene silencing 2)				
Plant Name	Accession #	MFE(kcal/mol)	Number of Loops	Number of Stems
<i>Arabidopsis thaliana</i>	AF239718	-190.49	33	30
<i>Oryza sativa</i>	NM_001049764	-203.99	36	32
7. SGS3 (Suppressor of gene silencing3)				
Plant Name	Accession #	MFE(kcal/mol)	Number of Loops	Number of Stems
<i>Arabidopsis thaliana</i>	AF239719	-214.36	31	30
<i>Oryza sativa</i>	NM_001072892	-223.67	35	30
8. eIF4E1 (Eukaryotic initiation factor 4E)				
Plant Name	Accession #	MFE(kcal/mol)	Number of Loops	Number of Stems
<i>Arabidopsis thaliana</i>	NM_117914	-173.96	36	30
<i>Oryza sativa</i>	NM_001052060	-222.00	32	30
9. IF(iso)4E (eukaryotic initiation factor(iso)4E)				
Plant Name	Accession #	MFE(kcal/mol)	Number of Loops	Number of Stems
<i>Arabidopsis thaliana</i>	AF538308	-179.58	35	30
<i>Oryza sativa</i>	NM_001070720	-220.24	35	32

10.{eIF4E1 (cum1)} (Cucumovirus multiplication 1)				
Plant Name	Accession #	MFE(kcal/mol)	Number of Loops	Number of Stems
<i>Arabidopsis thaliana</i>	AB107250	-171.06	38	31
<i>Oryza sativa</i>	NM_001071050	-308.45	29	26
11.{eIF4G (cum2)} (Cucumovirus multiplication 2)				
Plant Name	Accession #	MFE(kcal/mol)	Number of Loops	Number of Stems
<i>Arabidopsis thaliana</i>	AB107249	-165.27	35	32
<i>Oryza sativa</i>	NM_001066493	-205.35	31	28
12. RTM2 (Restricted TEV movement 2)				
Plant Name	Accession #	MFE(kcal/mol)	Number of Loops	Number of Stems
<i>Arabidopsis thaliana</i>	NM_120571	-149.43	35	29
<i>Oryza sativa</i>	NM_001052301	-244.15	23	19
13.HRT (Hyper sensitive response)				
Plant Name	Accession #	MFE(kcal/mol)	Number of Loops	Number of Stems
<i>Arabidopsis thaliana</i>	AF234174	-144.86	35	33
<i>Oryza sativa</i>	NM_001073080	-260.63	34	31
Housekeeping Genes of <i>Ath</i> & <i>Osa</i>				
1. Actin8 (ACT8)				
Plant Name	Accession #	MFE(kcal/mol)	Number of Loops	Number of Stems
<i>Arabidopsis thaliana</i>	NM_103814	-204.58	32	30
<i>Oryza sativa</i>	NM_001060954	-236.74	33	30
2. Polyubiquitin 10 (UBQ10)				
Plant Name	Accession #	MFE(kcal/mol)	Number of Loops	Number of Stems
<i>Arabidopsis thaliana</i>	NM_001084884	-188.72	33	30
<i>Oryza sativa</i>	NM_001052515	-223.03	36	33
3.Tubulin beta-2 (TUB2)				
Plant Name	Accession #	MFE(kcal/mol)	Number of Loops	Number of Stems
<i>Arabidopsis thaliana</i>	NM_125664	-215.55	34	32
<i>Oryza sativa</i>	NM_001057986	-261.61	34	31
4.Elongation factor-1 alpha (EF1-alpha)				
Plant Name	Accession #	MFE(kcal/mol)	Number of Loops	Number of Stems
<i>Arabidopsis thaliana</i>	NM_001125992	-202.72	30	29
<i>Oryza sativa</i>	NM_001055682	-249.07	28	27

maintenance of a database to store biological information at the beginning of the "genomic revolution", such as sequences of nucleotide and amino acid [29]. Various researchers use the bioinformatics tools for novel findings [30,31,32]. Codon usage bias has been extensively studied in Bacteria, as well as in Eukaryotes, there has been little work on patterns of codon bias in the third domain of life, the Archaea [33]. There are basically two approaches of measuring CUB. First, relative synonymous codon usage (RSCU) values can be calculated. RSCU is the observed number of codon occurrences divided by the number expected if synonymous codons were used uniformly. Second, the relative merits of different codons can be assessed from the viewpoint of translational efficiency [34]. The CUB analysis in this research was done by calculating relative synonymous codon usage (RSCU) values. The wonderful result was obtained in case of RTM1 gene in which all four codons do not use for alanine amino acid. Although synonymous codons encode the same amino acids, they are not used randomly and some are used more frequently than others Such as codon for Methionine is ATG and to more extant for Tryptophan amino acids the codon is TGG. Mukhopadhyay [35]. Reported similar CUB analyses for tissue specific and housekeeping genes in *A. thaliana* and *Oryza sativa*. The variation pattern of synonymous codon usage is well studied across the genes of an archaen, *Methanococcus maripalud* is which is most extensively studied archaen and is isolated from salt marsh sediment [36]. The use of alternative synonymous codon s is due to the divergence of coding sequences is well studied in closely related organism i.e., *Arabidopsis thaliana* and *Brassica rapa ssp. Pekinensis* [37]. The codon usage bias analysis under abiotic stress i.e., salt resistant genes in *Arabidopsis thaliana* and *Oryza sativa* is studied [38]. While not studied under biotic stress still, which is the main theme of this research. Recent Study is based on codon usage bias in biotic stress condition which is caused by viruses. A total of thirteen (13) virus resistant genes (RCY1, EDS1, EDS5, RTM1, LSP1, SGS2, SGS3, eIF4E, elf (iso) 4E, CUM1, CUM2, RTM2 and HRT) and four housekeeping genes (ACT8, TUB2, UBQ10 and EF1 α) were identified in *Arabidopsis thaliana* and *Oryza sativa*. The four housekeeping genes i.e., ACT8, TUB2, UBQ10 and EF1 α are also reported [39]. In the same way the genes like RCY1, EDS1, EDS5, RTM1, LSP1, SGS2, SGS3, eIF4E, elf (iso) 4E, CUM1, CUM2, RTM2 and HRT showing resistance to stress which is caused by virus infection in *Arabidopsis thaliana* and *Oryza sativa*, are also reported by [13,34,40,41,42,43,44,45,46].

Conclusion

The codon usage analysis was done for thirteen virus resistant and four housekeeping genes. In virus resistant genes out of thirteen 9 show similarity in the usage of codons by more than 50% and remaining 4 genes show similarity by less than 50% in *Arabidopsis thaliana* and *Oryza sativa*. Similarly out of four housekeeping genes only one gene show similarity by more than 50% and remaining 3 show similarities by less than 50%. The finding in this research is platform for engineering the virus resistant crops.

References

- Ikemura T (1985) Codon usage and tRNA content in unicellular and multicellular organisms. *J. Mol. Biol. Evol*2: 13-34.
- Bulmer M (1987) Coevolution of codon usage and transfer RNA abundance. *Nature* 325: 728-730.
- Akashi H (2001) Gene expression and molecular evolution. *Curr Opin Genet Dev*11: 660-666.
- Ermolaeva MD (2001) Synonymous Codon Usage in Bacteria. *Curr Issues Mol Bio*3: 91-97.
- Marais G, D Laurent (2000) Synonymous Codon Usage, Accuracy of Translation, and Gene Length in *Caenorhabditis elegans*. *J Mol Evol*52: 275-280.
- Balachandran S, Hurry VM, Kelley SE, Osmond CB, Robinson SA, Rohozinsk J, Seaton GGR, Sims DA (2006) Concepts of plant biotic stress. Some insights into the stress physiology of virus-infected plants, from the perspective of photosynthesis. *Plant Physiol*100: 203-213.
- Chisholm ST, Coaker G, Day B, Staskawicz BJ (2006) Host-microbe interactions: Shaping the evolution of the plant immune response. *Cell*124: 803-814.
- Jones JDG, Dangl JL (2006) The plant immune system. *Nature*444: 323-329.
- Eitas TK, Dangl JL (2010) NB-LRR proteins: pairs, pieces, perception, partners and pathways. *Curr Opin Plant Biol*13: 472-477.
- Kang BC, I. Yeam, MM Jahn (2005) Genetics of plant virus resistance. *Annu Rev Phytopatho*43: 581-621.
- Ponz F, Bruening G (1986) Mechanisms of resistance to plant viruses. *Annu Rev Phytopatho*24: 355-381.
- Fraile A, García-Arenal F (2010) The coevolution of plants and viruses: Resistance and pathogenicity. *Adv Virus Res*76: 1-32.
- Chisholm ST, Mahajan KS, Whitham SA, Yamamoto ML, Carrington JC (1999) Cloning of the *Arabidopsis* RTM1 gene, which controls restriction of long-distance movement of tobacco etch virus. *Proc Natl Acad Sci USA*97: 489-494.

14. Wang A, Krishnaswamy S (2012) Eukaryotic translation initiation factor 4E-mediated recessive resistance to plant viruses and its utility in crop improvement. *Mol Plant Pathol* 13:795-803.
15. Sekine SK, Hase SM, Kubota, Ichinose Y, Shah J, Kang HG, Klessig DF, Takahashi H (2008) High level expression of a virus resistance gene, RCY1, confers extreme resistance to Cucumber mosaic virus in *Arabidopsis thaliana*. *Mol Plant Microbe Interact.* 21: 1398-1407.
16. Peart JR, Cook G, Feys BJ, Parker JE, Baulcombe DC, Lane C (2002) An EDS1 orthologue is required for N-mediated resistance against tobacco mosaic virus. *The Plant J* 29: 569-579.
17. Ishihara T, Sekine K, Hase S, Kanayama Y, Seo S, Ohashi Y, Kusano T, Shibata D, Shah J, Takahashi H (2008) Overexpression of the EDS5 gene enhances resistance to viruses. *Plant Biol* 451-461. ISSN1435-8603.
18. Chisholm ST, Mahajan SK, Whitham SA, Yamamoto ML, Carrington JC (1999) Cloning of the *Arabidopsis* RTM1 gene, which controls restriction of long-distance movement of Tobacco etch virus. *Proc Natl Acad Sci* 97: 489-494.
19. Lellis AD, KD Kasschau, SA Whitham, JC Carrington (2002) Loss-of-Susceptibility Mutants of reveal an Essential Role for eIF (iso) 4E during Potyvirus Infection. *Curr Biol* 12: 1046-1051.
20. Mourrain P, Elmayan CB, Feuerbach F, Godon C, Morel J, Jouette D, Lacombe A, Nikic S, Picault N, Rémoúé K, Sanial M, Vo1 T, Vaucheret H (2000) *Arabidopsis* SGS2 AND SGS3 genes are required for Posttranscriptional Gene Silencing and Natural Virus Resistance. *Cell* 101: 533-542.
21. Sato M, Nakahara K, Yoshii M, Ishikawa M, Uyeda I (2005) Selective involvement of members of the eukaryotic initiation factor 4E family in the infection of by potyviruses. *FEBS Lett* 579: 1167-1171.
22. Duprat A, C Caranta, F Revers, B Menand, KS Browning, C Robaglia (2002) The *Arabidopsis* eukaryotic initiation factor (iso) 4E is dispensable for plant growth but required for susceptibility to potyviruses. *The Plant J* 32: 927-934.
23. Yoshii M, M Nishikiori, K Tomita, N Yoshioka, R Kozuka, S Naito, M Ishikawa (2004) The *Arabidopsis* cucumovirus multiplication 1 and 2 loci encode translation initiation factors 4E and 4G. *J of VIRO* 78: 6102-6111.
24. Whitham SA, Anderberg RJ, Chisholm ST, Carrington JC (2000) *Arabidopsis* RTM2 gene is necessary for specific restriction of tobacco etch virus and encodes an unusual small heat shock-like protein. *Plant Cell* 12: 569-582.
25. Cooley MB, Pathirana S, Wu HJ, Kachroo P, Klessig DF (2000) Members of the Arabidopsis HRT/RPP8 family of resistance genes confer resistance to both viral and oomycetepahogens. *Plant Cell* 12: 663-676.
26. Lilly ST, Drummond RSM, Pearson MN, MacDiarmid RM (2011) Identification and Validation of Reference Genes for Normalization of Transcripts from Virus-Infected *Arabidopsis thaliana*. *MPMI* 3:294-304.
27. Krieger M, Scott MP, Matsudaira PT, Lodish HF, Darnell JE, Lawrence Z, Kaiser C, Berk A (2004) Structure of Nucleic Acids. Molecular cell biology. New York: W.H. Freeman & CO. ISBN 0-7167-4366-3.
28. Zuker M (2003) Mfold web server for nucleic acid folding and hybridization prediction. *Nucleic Acids Res* 31:3406-3415.
29. Attwood TK, Gisel AN, Eriksson E, Bongcam RE (2011) Concepts, Historical Milestones and the Central Place of Bioinformatics in Modern Biology: A European Perspective. ISBN 978-953-307-282-1, 1-38.
30. Collins FS, Lander ES, Rogers J, Waterson RH (2004) Finishing the euchromatic sequence of the human genome *Natu* 431: 931-945.
31. Azuaje FJ, Heymann M., Ternes AM, Wienecke-Baldacchino A, Struck D, Moes D, Schneider R (2012) Bioinformatics as a driver, not a passenger, of translational biomedical research: Perspectives from the 6th Benelux bioinformatics conference. *J of Clin Bioinfo*, 2:7 doi: 10.1186/2043-9113-2-7.
32. Sifakis EG, Lambrou GI, Prentza A, Vlahopoulos S, Koutsouris D, Stathopoulou FT, Chatziioannou AA (2011) Elucidating the identity of resistance mechanisms to prednisolone exposure in acute lymphoblastic leukemia cells through transcriptomic analysis: A computational approach. *J of Clin Bioinfo*, 1(1):36 [Epub ahead of print].
33. Karlin S, Mrazek J, Ma J, Brocchieri L (2005) Predicted highly expressed genes in archaeal genomes. *Proc Natl Acad Sci* 102: 7303-7308.
34. McInerney JO (1998) GCUA (General Codon Usage Analysis). *Bioinformatics* 14: 372-373.
35. Mukhopadhyay P, Basak S, Ghosh TC (2008) Differential selective constraints shaping codon usage pattern of housekeeping and tissue-specific homologous genes of rice and *Arabidopsis*. *DNARes* 15: 347-356.
36. Emery LR, Sharp PM (2012) Impact of translational selection on codon usage Bias in the archaeon *Methanococcus maripaludis*. *Biol Lett* 7: 131-135.
37. Tiffin P, Hahn MW (2002) Coding sequence divergence between two closely related plant species: *Arabidopsis thaliana* and *Brassicarapasspekinensis*. *J Mol Evol* 54: 746-753.

38. Barozai KYM, Kakar GA, Din M (2012) The relationship between codon usage bias and salt resistant genes in *Arabidopsis thaliana* and *Oryza sativa*. *Pure Appl Bio*1: 48-51.
39. Lilly, S.T., R.S Drummond, M.N Pearson, R.M MacDiarmid. 2011. Identification and Validation of Reference Genes for Normalization of Transcripts from Virus-Infected. *MolPlantMicrobeInteract*24: 294-304.
40. Sekine K, Nandi A, Ishihara T, Hase S, Ikegami M, Shah J, Takahashi H (2004) Enhanced Resistance to Cucumber mosaic virus in the ssi2 Mutant Is mediated via an SA-Independent Mechanism. *The American PhytopatholSoci*17: 623-632.
41. Peart JR, Cook G, Feys BJ, Parker JE, Baulcombe DC, Lane C (2002) An EDS1 orthologue is required for N-mediated resistance against tobacco mosaic virus. *The Plant J*29: 569-579.
42. Sato M, Nakahara K, Yoshii M, Ishikawa M, Uyeda I (2005) Selective involvement of members of the eukaryotic initiation factor 4E family in the infection of by potyviruses. *FEBS Lett*579: 1167-1171.
43. Mourrain P, Elmayan CB, Feuerbach F, Godon C, Morel J, Jouette D, Lacombe A, Nikic S, Picault N, Rémoúé K, Sanial M, Vo1T, Vaucheret H (2000) *Arabidopsis* SGS2 AND SGS2 genes are required for Posttranscriptional Gene Silencing and Natural Virus Resistance. *Cell*101: 533-542.
44. Yoshii M, Nishikiori M, Tomita K, Yoshioka N, Kozuka R, Naito S, Ishikawa M (2004) The *Arabidopsis* cucumovirus multiplication 1 and 2 loci encode translation initiation factors 4E and 4G. *J VIROL*78: 6102-6111.
45. Whitham SA, Anderberg RJ, ChisholmST, Carrington JC (2000) *Arabidopsis* RTM2 gene is necessary for specific restriction of tobacco etch virus and encodes an unusual small heat shock-like protein. *Plant Cell*12: 569-582.
46. Cooley MB, Pathirana S, Wu HJ, Kachroo P, Klessig DF (2000) Members of the *Arabidopsis* HRT/RPP8 family of resistance genes confer resistance to both viral and oomycete pathogens. *Plant Cell*12: 663-676.