

Characterization Analysis of Codon Usage of Glutenin Subunits Genes in Wheat

Xingqin Cao¹, Yanping Ren², Jianbing Qin^{3*}

¹College of Computer Science, Yangtze University, Jingzhou Hubei

²College of Agriculture, Xinjiang Agricultural University, Urumqi Xinjiang

³College of Life Science, Yangtze University, Jingzhou Hubei

Email: *757593664@qq.com

Received: Mar. 7th, 2016; accepted: Mar. 21st, 2016; published: Mar. 24th, 2016

Copyright © 2016 by authors and Hans Publishers Inc.

This work is licensed under the Creative Commons Attribution International License (CC BY).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

Abstract

The codon characters of 13 high molecular weight glutenin subunits genes from wheat were analyzed by software CodonW1.4.2 and SPSS16.0. The results showed that ENC of HMW-GS genes ranges between 40 and 50, but 1Dy10 gene' is less than 40; GC3 contents are between 0.38 and 0.45; 1Bx14 gene prefers to use A or T at the end of codon, and the remaining 12 genes prefer to use G or C ending codons; the Euclidean distance coefficient is minimum between the alleles. The results indicated that the wheat HMW-GS genes have strong codon usage bias; the codon usage is similar; there are bias codons for each amino acid. Considering the codon usage differences, except 1By9 and 1Dy10, the difference of codon use frequency is minimum between x- or y-type subunit genes of glutenin in each chromosome group. The results have certain significance in the efficient use of HMW-GS genetic resources and the cloning of new LMW-GS genes.

Keywords

Wheat, HMW-GS, Codon Bias

小麦麦谷蛋白亚基基因密码子用法特性分析

曹兴芹¹, 任燕萍², 覃建兵^{3*}

¹长江大学计算机学院, 湖北 荆州

²新疆农业大学农学院, 新疆 乌鲁木齐

³长江大学生命科学学院, 湖北 荆州

*通讯作者。

Email: *757593664@qq.com

收稿日期: 2016年3月7日; 录用日期: 2016年3月21日; 发布日期: 2016年3月24日

摘要

该文运用CodonW1.4.2和SPSS16.0软件对13个来源于小麦麦谷蛋白亚基基因的密码子用法特性进行分析。结果显示,除1Dy10基因ENC值小于40外,其余亚基基因ENC值均介于40~50之间;GC3含量均在0.38~0.45之间;1Bx14基因偏向于使用A或T结尾的密码子,其余12个亚基基因偏向于使用G或C结尾的密码子;1Dy10与其余12个亚基基因间距离系数较大,然而各组染色体麦谷蛋白x-型或y-型亚基基因间距离系数较小。结果表明,小麦麦谷蛋白亚基基因具有较强的密码子使用偏爱性,各基因密码子用法上具有相似性,每一种氨基酸均存在1~2个偏爱密码子。在密码子用法差异上表现为,除1By9和1Dy10之外,无论是A染色体组、B染色体组,还是D染色体组,各组染色体上麦谷蛋白x-型或y-型亚基基因间密码子使用频率差异最小。这一分析结果对小麦HMW-GS功能基因资源的高效利用和LMW-GS新基因的克隆具有一定的指导意义。

关键词

小麦, 麦谷蛋白亚基, 密码子偏爱性

1. 引言

某一物种或某一类基因普遍存在同义密码子非均衡使用的现象,即密码子偏好性,其中通常倾向于使用的一种或几种特定的同义密码子被称为最优密码子[1]。大肠杆菌[2]、细菌[3]、古菌[4]、拟南芥[5]、玉米[6]、水稻[7]、籼稻[8]、杨树[9]、鱼[10]、牦牛[11]等不同物种的某一类功能基因在密码子使用上存在着明显的偏性,而且某些特定基因[12]密码子使用频率与其功能在一定的范围内呈现正相关趋势。然而,有关小麦(*Triticum aestivum*)麦谷蛋白亚基基因密码子用法特征分析还少见报道,也值得进一步探究。

小麦是国内外最重要的粮食作物之一。我国面包小麦严重缺乏,面包粉消费几乎完全依赖进口。目前,应用麦谷蛋白亚基组成和表达去改良小麦面包烘烤品质研究仍是该领域研究的焦点问题[13][14]。因此,深入探究小麦麦谷蛋白亚基基因密码子使用特性、不同小麦麦谷蛋白亚基基因密码子用法的差异,为进一步深入开展小麦高分子量麦谷蛋白亚基(high molecular weight glutenin subunit, HMW-GS)基因功能研究提供基础数据,同时也为相对滞后的小麦低分子量麦谷蛋白亚基(low molecular weight glutenin subunit, LMW-GS)基因的克隆、表达、结构和功能研究提供科学的理论数据参考。

因此,该文拟以NCBI数据库中提供的小麦麦谷蛋白亚基基因完整的全长编码区序列(coding DNA sequence, 简称CDS)为对象,运用CodonW1.4.2软件分析麦谷蛋白亚基基因的密码子使用特性,并结合SPSS16.0软件使用聚类分析的方法比较不同麦谷蛋白亚基基因密码子用法的差异,希望能够为小麦HMW-GS功能基因资源的高效利用和LMW-GS新基因的克隆提供参考。

2. 数据和分析方法

2.1. 小麦麦谷蛋白亚基基因密码子使用特性分析

该文下载的小麦麦谷蛋白亚基基因完整的CDS序列来源于GenBank,其在NCBI数据库(<http://www.ncbi.nlm.nih.gov/>)的登录号见表1。

Table 1. The accession number of HMW-GS genes in wheat
表 1. 小麦 HMW-GS 基因序列登录号

序列登录号	亚基
DQ533690.1	1Ax1
EF055262.1	1Ax2*
AY367771.1	1Bx14
AY553933.1	1Bx23
EF381742.1	1By9
DQ086215.1	1By15
EF540765.1	1By16
GQ241722.1	1Dx1.5
AY517724.2	1Dx2
HM050419.1	1Dx5
DQ211819.1	1Dy10
EU528008.1	1Dy11
EU266533.1	1Dy12

运用密码子分析软件 CodonW1.4.2, 计算以下衡量同义密码子使用偏性的主要参数: 有效密码子数(ENC)、全长基因 GC 含量(GC)、密码子第三位 GC 含量(GC3)、同义密码子相对使用度(RSCU)。

2.2. 基于密码子使用偏好性的聚类

利用 SPSS16.0 软件对 13 个来源于小麦麦谷蛋白亚基基因进行基于密码子使用偏性的聚类分析。在对基因密码子使用概率分析时, 将每一条基因作为一个对象, 同义密码子的 RSCU 值作为变量, 采用 59 个同义密码子的 RSCU 值对其密码子使用偏性进行分析。基因间的距离规定为同义密码子相对使用度的欧拉平方距离[15]。构建距离系数矩阵, 并进行基于该系数的聚类分析。

3. 结果与分析

3.1. 麦谷蛋白亚基基因同义密码子使用偏性分析

来源于小麦的 13 个麦谷蛋白亚基基因同义密码子使用偏性主要参数 ENC、GC、GC3 计算结果见表 2。ENC 值能反映密码子家族中同义密码子偏爱程度。其范围在 20~61 之间, 越靠近 20 偏性越强。研究结果显示, 1Ax1、1Ax2*、1Bx14、1Bx23、1By9、1By15、1By16、1Dx1.5、1Dx2、1Dx5、1Dy11、1Dy12* 基因的 ENC 值均介于 40~50 之间, 表明它们的密码子使用偏爱性中等; 而 1Dy10 基因的 ENC 值为 36.76, 表明该基因具有较强的密码子使用偏爱性; 进一步分析小麦的麦谷蛋白亚基基因 GC 含量发现, 不同麦谷蛋白亚基基因 GC 含量较高, 均在 0.5 以上; 从密码子第三位 GC 含量来看, 各亚基基因 GC3 含量在 0.38~0.45 之间, 没有太大差异。由此可见, 不同的小麦麦谷蛋白亚基基因的密码子用法上具有相似性。

3.2. 麦谷蛋白亚基基因同义密码子使用频率分析

运用 CodonW1.4.2 软件计算并分析不同的小麦麦谷蛋白亚基基因密码子使用情况, RSCU 值去除了氨基酸组成对密码子使用的影响, 且直观地反映了密码子使用的偏爱性。如某一密码子的 RSCU 值等于 1, 则表明该密码子的使用没有偏爱; RSCU 值大于 1, 表明该密码子的使用频率相对较高; RSCU 值小

Table 2. The ENC values and content of GC for glutenin subunits gene in wheat
表 2. 不同麦谷蛋白亚基基因的有效密码子数和 GC 含量

亚基	有效密码子数	全长 CDS 的 GC 含量	密码子第三位 GC 含量
1Ax1	42.23	0.547	0.405
1Ax2*	44.13	0.550	0.408
1Bx14	43.68	0.552	0.400
1Bx23	46.45	0.552	0.398
1By9	48.48	0.548	0.495
1By15	42.47	0.546	0.396
1By16	43.60	0.542	0.388
1Dx1.5	46.42	0.572	0.429
1Dx2	45.88	0.573	0.427
1Dx5	46.46	0.573	0.426
1Dy10	36.76	0.543	0.446
1Dy11	46.30	0.558	0.418
1Dy12*	44.43	0.555	0.412

于 1, 表明该密码子的使用频率低。研究结果显示, 1Ax1 基因有 23 个 RSCU 值大于 1 的密码子, 1Ax2* 有 22 个, 1Bx23 有 24 个, 1By9 有 28 个, 1By15 有 27 个, 1By16 有 27 个, 1Dx1.5 有 23 个, 1Dx2 有 24 个, 1Dx5 有 23 个, 1Dy10 有 21 个, 1Dy11 有 25 个, 1Dy12* 有 24 个, 且这些密码子绝大部分以 G 或 C 碱基结尾; 而 1Bx14 基因有 23 个 RSCU 值大于 1 的密码子, 且大部分以 A 或 T 碱基结尾(表 3)。由此可见, 小麦麦谷蛋白亚基基因偏爱 G 或 C 结尾的密码子。

另外, 根据同义密码子相对使用度可以看出, 在小麦麦谷蛋白亚基中, 编码 Ala 的偏爱密码子为 GCA, 编码 Cys 的偏爱密码子为 TGC, 编码 Asp 的偏爱密码子为 GAC, 编码 Glu 的偏爱密码子为 GAG, 编码 Phe 的偏爱密码子为 TTT, 编码 Gly 的偏爱密码子为 GGA/GGG, 编码 His 的偏爱密码子为 GAC, 编码 Ile 的偏爱密码子为 ATA, 编码 Lys 的偏爱密码子为 AAG, 编码 Leu 的偏爱密码子为 CTA/CTC, 编码 Asn 的偏爱密码子为 AAC, 编码 Pro 的偏爱密码子为 CCA, 编码 Gln 的偏爱密码子为 CAA, 编码 Arg 的偏爱密码子为 AGG, 编码 Ser 的偏爱密码子为 TCA/TCT, 编码 Thr 的偏爱密码子为 ACT, 编码 Val 的偏爱密码子为 GTC/GTG, 编码 Tyr 的偏爱密码子为 TCA。由此可见, 不同的小麦麦谷蛋白亚基基因表达, 每一种氨基酸均存在 1~2 个偏爱密码子。

3.3. 不同麦谷蛋白亚基基因密码子使用频率的差异

基于表 3 中各密码子的 RSCU 值, 利用 SPSS 计算不同麦谷蛋白亚基基因之间密码子使用偏性的欧式距离系数(表 4), 并进行基于该系数的聚类分析(图 1)。各基因间的距离系数越大, 则表达它们在密码子使用上的差异越大。由表 4 可以看出, 1Dy10 与 1Dx1.5、1Dx2 和 1Dx5 基因间距离系数最大, 和 1Ax1、1Ax2*、1Bx14、1Bx23、1By9、1Dy11 和 1Dy12* 基因间距离系数次之, 而各组染色体上麦谷蛋白 x-型或 y-型亚基基因间距离系数较小。研究表明, 1Dy10 与其他麦谷蛋白亚基基因的密码子使用差异较大, 而各组染色体上麦谷蛋白 x-型或 y-型亚基基因之间密码子使用差异较小。

基于距离系数的聚类分析可知, 1Dx1.5、1Dx2 和 1Dx5 基因之间密码子用法差异相对较小, 成为一

Table 3. The relative synonymous codon usage of glutenin subunits genes in wheat
表 3. 不同的小麦麦谷蛋白亚基基因同义密码子的相对使用度

氨基酸	密码子	1Ax1	1Ax2*	1Bx14	1Bx23	1By9	1By15	1By16	1Dx1.5	1Dx2	1Dx5	1Dy10	1Dy11	1Dy12*
A(Ala)	GCA	1.28	1.33	1.75	1.73	0.92	1.13	0.62	2.00	2.07	2.07	2.99	0.93	0.90
	GCC	0.64	0.67	0.75	0.67	1.54	0.50	0.62	0.57	0.55	0.55	0.69	0.93	0.77
	GCG	1.44	1.17	1.00	1.07	0.31	0.38	0.46	0.71	0.69	0.69	0.05	0.67	0.65
	GCT	0.64	0.83	0.50	0.53	1.23	2.00	2.31	0.71	0.69	0.69	0.27	1.47	1.68
C(Cys)	TGC	1.50	1.20	1.00	1.50	1.60	1.43	1.43	1.50	1.50	1.50	0.00	1.43	1.43
	TGT	0.50	0.80	1.00	0.50	0.40	0.57	0.57	0.50	0.50	0.50	2.00	0.57	0.57
D(Asp)	GAC	1.33	1.33	1.33	1.60	1.60	1.50	1.50	2.00	2.00	2.00	0.00	1.50	1.50
	GAT	0.67	0.67	0.67	0.40	0.40	0.50	0.50	0.00	0.00	0.00	0.00	0.50	0.50
E(Glu)	GAA	0.61	0.75	0.29	0.43	0.60	0.59	0.71	0.25	0.24	0.50	1.00	0.71	0.59
	GAG	1.39	1.25	1.71	1.57	1.40	1.41	1.29	1.75	1.76	1.50	1.00	1.29	1.41
F(Phe)	TTC	1.00	1.00	0.80	1.00	0.67	0.67	0.67	0.67	0.67	0.67	1.71	1.33	0.67
	TTT	1.00	1.00	1.20	1.00	1.33	1.33	1.33	1.33	1.33	1.33	0.29	0.67	1.33
G(Gly)	GGA	1.80	1.73	2.22	2.20	1.30	1.76	1.76	1.65	1.64	1.60	1.41	1.77	1.78
	GGC	0.28	0.34	0.15	0.14	0.80	0.22	0.21	0.17	0.17	0.17	0.94	0.22	0.21
	GGG	1.55	1.62	1.55	1.49	1.60	1.95	1.97	1.80	1.82	1.85	1.41	1.91	1.91
	GGT	0.37	0.31	0.09	0.11	0.30	0.06	0.06	0.37	0.37	0.38	0.24	0.11	0.10
H(His)	CAC	1.20	1.60	0.00	0.00	2.00	1.09	1.09	1.00	1.00	1.00	2.00	1.38	1.47
	CAT	0.80	0.40	2.00	2.00	0.00	0.91	0.91	1.00	1.00	1.00	0.00	0.62	0.53
I(Ile)	ATA	2.25	2.00	1.20	1.50	2.00	2.33	2.10	1.50	1.50	1.50	1.29	2.25	2.25
	ATC	0.00	0.00	0.60	0.75	0.50	0.33	0.60	0.75	0.75	0.75	1.71	0.75	0.75
	ATT	0.75	1.00	1.20	0.75	0.50	0.33	0.30	0.75	0.75	0.75	0.00	0.00	0.00
K(Lys)	AAA	0.00	0.00	0.29	0.33	0.60	0.57	0.75	0.57	0.57	0.57	0.33	0.89	0.89
	AAG	2.00	2.00	1.71	1.67	1.40	1.43	1.25	1.43	1.43	1.43	1.67	1.11	1.11
L(Leu)	CTA	1.83	1.86	1.38	1.38	0.67	1.20	1.76	1.65	1.65	1.61	3.30	1.45	1.66
	CTC	0.78	0.71	1.85	1.62	2.67	1.40	1.24	1.50	1.50	1.46	1.20	1.66	1.66
	CTG	0.39	0.57	2.08	1.85	1.33	2.60	2.47	1.50	1.50	1.61	0.90	1.86	1.86
	CTT	0.39	0.57	0.00	0.00	0.67	0.20	0.18	0.00	0.00	0.00	0.00	0.41	0.21
	TTA	1.04	0.57	0.00	0.23	0.00	0.20	0.00	0.15	0.15	0.15	0.30	0.00	0.00
	TTG	1.57	0.71	0.69	0.92	0.67	0.40	0.35	1.20	1.20	1.17	0.30	0.62	0.62
N(Asn)	AAC	0.00	0.00	0.00	0.00	1.33	2.00	2.00	0.00	0.00	0.00	1.78	0.00	0.00
	AAT	0.00	0.00	0.00	0.00	0.67	0.00	0.00	0.00	0.00	0.00	0.22	0.00	0.00
P(Pro)	CCA	2.85	2.88	2.88	2.88	2.96	3.44	3.43	2.91	2.92	2.93	1.81	3.27	3.38
	CCC	0.16	0.17	0.17	0.13	0.17	0.17	0.17	0.15	0.15	0.15	1.42	0.12	0.11
	CCG	0.95	0.92	0.83	0.86	0.35	0.28	0.29	0.91	0.90	0.89	0.26	0.36	0.28
	CCT	0.04	0.04	0.13	0.13	0.52	0.11	0.11	0.04	0.04	0.04	0.52	0.24	0.23

续表

Q(Gln)	CAA	1.32	1.31	1.48	1.45	1.37	1.43	1.48	1.35	1.37	1.36	1.44	1.41	1.40
	CAG	0.68	0.69	0.52	0.55	0.63	0.57	0.52	0.65	0.63	0.64	0.56	0.59	0.60
R(Arg)	AGA	2.45	2.45	0.96	1.43	0.55	0.63	1.26	0.00	0.00	0.00	0.47	0.43	0.46
	AGG	1.91	1.09	2.16	1.71	1.09	2.21	1.58	0.67	0.55	0.50	5.29	1.29	1.85
	CGA	0.82	1.36	0.72	0.86	1.09	0.95	0.95	0.67	1.09	1.50	0.08	1.29	0.92
	CGC	0.27	0.27	1.20	1.14	1.09	0.63	0.32	1.33	1.09	1.00	0.00	0.86	0.46
	CGG	0.55	0.82	0.96	0.86	2.18	1.58	1.58	2.67	2.73	2.50	0.16	2.14	2.31
	CGT	0.00	0.00	0.00	0.00	0.00	0.00	0.32	0.67	0.55	0.05	0.00	1.00	0.00
	AGC	0.54	0.40	0.55	0.52	1.89	0.84	0.75	0.73	0.73	0.75	1.20	1.35	1.26
S(Ser)	AGT	0.21	0.20	0.09	0.09	0.00	0.12	0.13	0.12	0.12	0.13	0.40	0.00	0.00
	TCA	2.25	2.40	2.40	2.43	0.63	1.32	1.38	1.59	1.59	1.50	1.00	0.90	0.84
	TCC	0.00	0.00	0.18	0.35	0.32	0.12	0.38	0.00	0.00	0.00	1.80	0.30	0.28
	TCG	0.32	0.60	0.65	0.70	0.32	0.36	0.38	0.73	0.73	0.75	0.00	0.30	0.28
	TCT	2.68	2.40	2.12	1.91	2.84	3.28	3.00	2.82	2.82	2.88	1.60	3.15	3.35
	ACA	0.43	0.44	0.14	0.00	0.71	1.44	1.71	0.32	0.32	0.32	3.19	1.17	1.12
T(Thr)	ACC	0.29	0.30	0.57	0.55	1.18	0.48	0.43	0.32	0.32	0.32	0.55	0.50	0.48
	ACG	0.14	0.15	0.29	0.28	0.47	0.32	0.29	0.16	0.16	0.16	0.04	0.33	0.32
	ACT	3.14	3.11	3.00	3.17	1.65	1.76	1.57	3.20	3.20	3.20	0.21	2.00	2.08
V(Val)	GTA	0.71	1.05	0.53	0.53	1.07	1.14	1.14	0.57	0.57	0.57	1.78	0.84	0.84
	GTC	1.18	1.26	1.07	1.07	1.60	1.14	1.14	1.71	1.71	1.71	1.33	1.26	1.26
	GTG	0.94	0.84	1.87	1.87	1.07	1.33	1.33	1.33	1.33	1.33	0.67	1.47	1.47
	GTT	1.18	0.84	0.53	0.53	0.27	0.38	0.38	0.38	0.38	0.38	0.22	0.42	0.42
Y(Tyr)	TAC	1.83	1.76	1.89	1.89	1.33	1.55	1.59	1.91	1.91	1.91	1.00	1.63	1.60
	TAT	0.17	0.24	0.11	0.11	0.67	0.45	0.41	0.09	0.09	0.09	1.00	0.38	0.40

注：大于 1 的 RSCU 值用黑体字表示。

亚类；1Bx14 和 1Bx23 成为一亚类；1Ax1 和 1Ax2*成为一亚类；1By15 和 1By16 成为一亚类；1Dy11 和 1Dy12*成为一亚类；最后再与密码子用法差异最大的 1Dy10 成为一类。这一聚类分析结果同样表明，除 1By9 和 1Dy10 之外，无论是 A 染色体组、B 染色体组，还是 D 染色体组，各组染色体上麦谷蛋白 x-型或 y-型亚基因间密码子使用频率差异最小。

4. 讨论

该文对 13 个小麦麦谷蛋白亚基因密码子的用法分析发现，麦谷蛋白亚基因的 GC 含量高。进一步分析发现，该类基因每个氨基酸在同义密码子的使用上都存在明显的偏爱性，大部分麦谷蛋白亚基因主要偏爱于使用 G 或 C 碱基结尾的密码子，少量使用 A 或 T 碱基结尾的密码子。该类基因与前人报道的单子叶植物基因组密码子偏爱特性趋势一致[16]。

密码子使用偏性产生的生物学基础目前还不是很清楚。近几年研究中，通过对模式生物的密码子用法研究，表明同义密码子使用偏性主要由突变偏性、翻译选择或两者共同作用产生[17]。不同物种或同一物种的不同基因对密码子的用法特性有所不同，这种密码子使用偏爱性差异大小与物种的亲缘关系远近有关，亲缘关系越远，密码子偏爱性差异越大，越近则差异越小[18]。该文研究结果表明，在小麦麦谷蛋

Table 4. Coefficient of squared Euclidean distance of codon usage bias between samples of glutenin subunits gene
表 4. 各麦谷蛋白亚基基因间相对密码子使用度的欧氏距离系数

	1Ax1	1Ax2*	1Bx14	1Bx23	1By9	1By15	1By16	1Dx1.5	1Dx2	1Dx5	1Dy10	1Dy11	1Dy12*
1Ax1	—												
1Ax2*	1.822	—											
1Bx14	4.227	4.312	—										
1Bx23	3.717	3.894	1.378	—									
1By9	5.797	5.419	5.646	5.651	—								
1By15	5.216	5.077	4.663	4.788	3.825	—							
1By16	5.243	4.934	5.001	4.980	4.086	1.506	—						
1Dx1.5	4.762	4.543	3.754	3.644	4.505	4.580	4.837	—					
1Dx2	4.782	4.488	3.837	3.699	4.506	4.600	4.842	0.524	—				
1Dx5	4.712	4.345	3.821	3.643	4.429	4.486	4.753	1.186	0.805	—			
1Dy10	8.577	8.625	8.588	8.903	8.385	7.696	7.804	9.268	9.335	9.307	—		
1Dy11	4.970	4.688	4.655	4.563	3.296	3.188	3.194	3.476	3.477	3.464	8.195	—	
1Dy12*	4.881	4.726	4.610	4.646	3.300	2.747	2.918	3.590	3.589	3.503	8.031	1.672	—

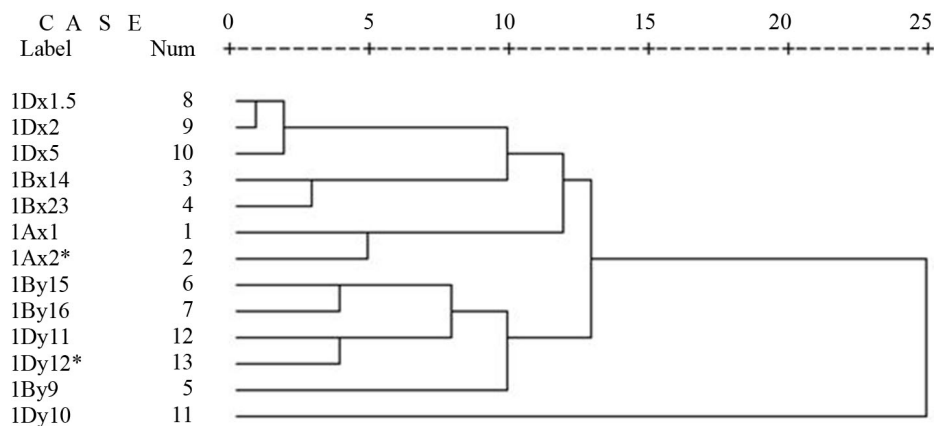


Figure 1. Cluster analysis dendrogram of RSCU for glutenin subunits genes

图 1. 基于麦谷蛋白亚基基因相对密码子使用度的聚类图

白亚基基因之间，1Dy10 与其他麦谷蛋白亚基基因的密码子使用差异较大，然而无论是 A 染色体组、B 染色体组，还是 D 染色体组，各组染色体上麦谷蛋白 x-型或 y-型亚基基因之间密码子使用差异较小；除 1By9 和 1Dy10 之外，各组染色体麦谷蛋白 x-型或 y-型亚基基因间密码子使用频率差异最小。

因此，基于目前编码区密码子使用偏好性和使用信息的聚类分析，可为进一步深入研究小麦 HMW-GS 基因功能提供基础数据，同时也为相对滞后的小麦 LMW-GS 基因的克隆、表达、结构和功能研究提供科学的理论数据参考。

参考文献 (References)

- [1] 郭秀丽, 王玉, 杨路成, 等. 茶树 CBF1 基因密码子使用特性分析[J]. 遗传, 2012, 34(12): 1614-1623.

- [2] 王侃, 刘次全, 曹槐, 等. 大肠杆菌 mRNA 编码区长度、形成二级结构倾向与密码子偏好性的关系[J]. 微生物学报, 2006, 46(6): 895-899.
- [3] 钟东, 赵贵军, 刘宇虎, 等. 62 种细菌基因组中密码子碱基位置上碱基分布的差异性及相关性[J]. 中国微生态学杂志, 2003, 15(3): 134-136.
- [4] Wu, H.N., Da, Y., Wei, J.W., *et al.* (2007) Synonymous Codon Usage in *Methanosarcinamazei str. Goe1* and Other *Euryarchaeota* Microorganisms. *Journal of Southeast University*, **123**, 289-293.
- [5] 范三红, 郭蔼光, 单丽伟, 等. 拟南芥基因密码子偏爱性分析[J]. 生物化学与生物物理进展, 2003, 30(2): 221-225.
- [6] 赵耀, 刘汉梅, 顾勇, 等. 玉米 waxy 基因密码子偏好性分析[J]. 玉米科学, 2008, 16(2): 16-21.
- [7] 张晓峰, 薛庆中. 水稻和拟南芥 NBS-LRR 基因家族同义密码子使用偏好的比较[J]. 作物学报, 2005, 31(5): 596-602.
- [8] 刘庆坡, 谭军, 薛庆中. 籼稻品种 93-11 同义密码子的使用偏性[J]. 遗传学报, 2003, 30(4): 335-340.
- [9] 单文娟, 周猛, 卫巍, 等. 杨树部分基因起始密码子 AUG 侧翼序列特征分析[J]. 南京林业大学学报(自然科学版), 2008, 32(5): 61-64.
- [10] 龙华, 王祖荣. 15 种鱼转铁蛋白 cDNA 序列密码子使用比较分析[J]. 南方水产, 2005, 1(1): 6-10.
- [11] 赵胜, 张琴, 廖伟璇, 等. 41 条牦牛 CDS 序列的密码子偏好性分析[J]. 西南民族大学学报, 2005, 31(5): 755-760.
- [12] 刘华贵, 王小花, 刘玉方, 等. 五个鸡群中催乳素基因密码子使用频率与产蛋量的关系分析[J]. 生物化学与生物物理进展, 2007, 34(10): 1101-1106.
- [13] 武茹, 高德荣, 别同德, 等. 转基因小麦 1Dx5 基因沉默的遗传表达和品质效应分析[J]. 麦类作物学报, 2010, 30(6): 991-996.
- [14] 祝长青, 赵惠新, 李艳红, 等. 转基因小麦后代外源目的基因的遗传分析[J]. 干旱区研究, 2008, 25(6): 808-811.
- [15] Das, S., Paul, S. and Dutta, C. (2006) Synonymous Codon Usage in Adenoviruses: Influence of Mutation, Selection and Protein Hydropathy. *Virus Research*, **117**, 227-236. <http://dx.doi.org/10.1016/j.virusres.2005.10.007>
- [16] 刘汉梅, 赵耀, 顾勇, 等. 几种植物 waxy 基因的密码子用法特性分析[J]. 核农学报, 2010, 24(3): 476-481.
- [17] Liu, Q., Feng, Y., Zhao, X., *et al.* (2004) Synonymous Codon Usage Bias in *Oryza sativa*. *Plant Science*, **167**, 101-105. <http://dx.doi.org/10.1016/j.plantsci.2004.03.003>
- [18] 吴宪明, 吴松锋, 任大明. 密码子偏性的分析方法及相关研究进展[J]. 遗传, 2007, 29(4): 420-426.