

Analysis of Soybean Protein Content for Environment Stability Based on GGE Biplot

Xue Han¹, Huidong Qi², Zhaoming Qi^{2*}, Qingshan Chen²

¹Beidahuang Kenfeng Seed Company, Harbin Heilongjiang

²Northeast Agricultural University, Harbin Heilongjiang

Email: qizhaoming1860@126.com, qshchen@126.com

Received: Jul. 4th, 2018; accepted: Jul. 19th, 2018; published: Jul. 26th, 2018

Abstract

To increase the evaluation accuracy of stability and adaptability of soybean protein content in multiple environments, GGE-biplot method was used to analyze and evaluate the recombinant inbred lines (RILs) dataset including 147 F₁₀₋₁₉ RILs and their parents Charleston (♀) and Dongnong 594 (♂) planted in Harbin from 2002 to 2011 year. The results showed that the first principal component explained 37.7% effect, the second principal components explained 15.3% effect, and PC1 and PC2 were explained totally 53.0% on interaction effects. Ten environments could be divided into four groups, 2003HRB was in the first group, and Line No. 80 was the best inside the group; 2002HRB, 2004HRB, 2006HRB and 2008HRB were in the second group, and Line No. 24 was the best inside the group; 2011HRB was in the third group, and Line No. 10 was the best inside the group. 2007HRB, 2009HRB and 2010HRB were in the second group, and Line No. 23 was the best inside the group. Based on discrimination ability and environmental representativeness, 2008HRB was the best representative environment and 2002HRB was the best discriminating ability environment. Based 10 years data in Harbin environment, environmental factors impacted on the soybean protein content are sequenced as: Sunshine > precipitation > average temperature. The results were important implications for guiding soybean cultivation in multiple environments and evaluation of the suitable environmental factors.

Keywords

Soybean, Protein Content, Environment Factor, GGE Biplot

基于GGE双标图的大豆蛋白质多环境的稳定性分析

韩雪¹, 齐慧冬², 齐照明^{2*}, 陈庆山²

*通讯作者。

¹北大荒垦丰种业有限公司, 黑龙江 哈尔滨

²东北农业大学, 黑龙江 哈尔滨

Email: qizhaoming1860@126.com, qshchen@126.com

收稿日期: 2018年7月4日; 录用日期: 2018年7月19日; 发布日期: 2018年7月26日

摘要

为准确评价大豆蛋白质含量在不同环境的稳定性和适应性, 本研究应用GGE双标图对2002年到2011年种植于哈尔滨试验点的147个F₁₀₋₁₉代重组自交系材料及其亲本Charleston (♀)和东农594 (♂)的10个年份(环境)蛋白数据进行分析 and 评价。结果表明, 第1主成了解释了37.7%的效应, 第2主成了解释了15.3%的效应, 第1主成分和第2主成分共解释G与GE互作效应的53.0%。10个环境被分为4个区域, 第一个区域主要有2003HRB环境, 株系80的蛋白质含量表现最好; 第二个区域有2002HRB、2004HRB、2006HRB和2008HRB共4个环境, 株系24的蛋白质含量表现最好; 第三个区域主要有2011HRB, 株系10的蛋白质含量表现最好。第四个区域有2007HRB、2009HRB和2010HRB, 株系23位于顶点, 蛋白质含量表现最好。通过环境的代表性和鉴别能力分析, 得出2008HRB最具有代表性, 2002HRB鉴别能力最好。在连续10年哈尔滨环境下, 环境因子对蛋白质含量的影响顺序为: 日照时数 > 降水量 > 平均温度。该结果对多环境下指导大豆栽培及地区适宜气象因子的评估具有重要意义。

关键词

大豆, 蛋白质含量, 环境因子, GGE双标图

Copyright © 2018 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

1. 引言

大豆发源地于亚洲, 是重要的粮食作物之一, 大豆的蛋白质提供了人体蛋白的主要来源。含有 20 多种自然界普遍存在的氨基酸, 营养价值比较高, 与动物蛋白相比不含有胆固醇, 为很好的保健食品, 豆类蛋白还具有降血压, 降血脂, 减少心脑血管病以及促进营养吸收等功效[1]。大豆蛋白质含量性状为数量性状, 遗传机制比较复杂, 很容易受到如温度、光照和水分等外界环境条件影响, 环境因素的变化会导致数量性状基因的表达方式或表达程度发生变化, 进而影响蛋白质含量品种在各地点的不一致。因此, 将环境因子多元化有助于发掘真正影响大豆蛋白质含量的因素, 为大豆品质育种奠定了基础。

目前, 基于环境因子与表型的分析研究相对较少, 大多数都是研究 QTL 与环境的互作分析效应, 研究多采用 QTL 与环境互作的分析软件[2] [3], 但这些分析软件并不能直观地综合反映区域试验品种(系)的高产稳产性、试验点评价和品种生态区划分情况。1971 年, 双标图(Biplot)的概念被 Gabriel 等提出以来[4], 很多研究应用双标图进行了品种试验的有效区分[5] [6]。

近些年来, GGE-Biplot 双标图分析方法应用广泛。在小麦[7]、燕麦[8]、大豆[9]、小豆[10]、玉米[11]、棉花[12]和甘蔗[13]等作物都有研究。张勇等[14]对我国春麦区 10 个试点 20 个品种淀粉糊化特性的峰值粘度进行稳定性分析。罗俊[15]等对 5 个试点中的 7 个品种的甘蔗产量性状进行稳定性和适应性的分析。杨进文等[16]用 GGE 双标图及隶属函数综合分析山西小麦地方品种抗旱性。张春明等[17]也利用该分析

方法,对3个年份的10个品种19个试验点的小豆进行高产、稳产性和适应性分析。韩英鹏等在2014年利用GGE-Biplot方法分别对大豆的亚麻酸[18]和脂肪酸[19]性状表现进行分析。

本研究利用GGE双标图分析软件(GGE biplot)针对以Charleston和东农594为亲本的147个 F_{10-19} 重组自交系群体,分析其连续10年的蛋白质含量数据。揭示RIL群体147个株系在多环境下的品质性状表现及其稳定性,同时评价环境的区分力与代表性,能够为大豆栽培生产提供指导性依据。

2. 材料和方法

2.1. 试验材料

本研究以美国半矮秆品种Charleston(♀)和东农594(♂)构建的RIL群体繁殖得到的147个 F_{10-19} 代重组自交系为试验材料,其中Charleston是高油材料,其油分含量为22.7%;东农594是高蛋白材料,蛋白质含量为44.3%。

2.2. 材料种植

将Charleston(♀)和东农594(♂)及其147个 F_{10-19} 代重组自交系材料从2002年到2011年种植于黑龙江省农垦科研育种中心试验地(哈尔滨,126°38'E,45°45'N)。2002~2006年间种植两次重复,2007~2011年间种植三次重复,田间管理同常规大田。

2.3. 品质性状及环境因子测定

通过FOSS-1241红外线谷物分析仪对RIL群体147个 F_{10-19} 代重组自交系进行蛋白含量测定,每个株系随机取5株,保证其在安全含水量范围内测定蛋白质含量,取其均值作为该行的蛋白质含量,三次重复种植的均值作为该环境下蛋白质含量表型值。

平均温度、降水量和日照时数等主要环境因子是影响大豆蛋白含量的主要因素。《中国地面气候资料日值数据集数据格式》统计记录哈尔滨10年(2002年~2011年)的环境因子平均温度、降水量和日照时数等数据。

2.4. 分析方法

采用GGE-Biplot软件[5]进行大豆蛋白质含量位于多环境下的GGE-Biplot分析,利用以环境中心化的双标图,分析G与GE的效应,G:表示表型的主效应,GE表示:表型与环境互作的效应,所得的图示化结果用于评价重组自交系群体在试验环境上的表现。

GGE双标图采用平均环境坐标(Average environment coordinate, ACE)法能够使双向数据表中的信息清晰、全面的显示出来。AEC是用圆圈表示的环境平均值,AEC的PC1和PC2得分分别等于所有指标的PC1和PC2的平均得分。AEA(Average-environment axis)即AEC的横轴,通过原点和AEC,AEC的纵轴就是通过原点垂直于AEA的直线。GGE双标图上同时给出蛋白质性状和环境图标,可更直观地进行蛋白质性状品质的分析,AEA表示大豆的平均含油量,箭头方向为正,各株系在AEA上的投影越靠右,其蛋白质品质越高,纵轴代表了株系的稳定性,箭头方向株系的稳定性差,AEA上垂线的长短显示稳定性的大小,长短越短,稳定性越好。

将GGE双标图结果中显示出的各环境最优点点用直线连接起来,形成一个多边形囊括所有环境。以原点为顶点向各边出发,作各边垂线将图分为若干区域,每个环境自然落于某个区域内,靠近原点的株系对环境变化迟钝,而“顶点”所代表的株系是该区域内在每个环境都表现最优的株系。并解析大豆蛋白质性状品质在哈尔滨10年环境受主要环境因子(平均温度、降水量和日照时数)的影响,找出影响大

豆蛋白质含量的最优环境因子。

3. 结果与分析

3.1. 环境因子统计分析

对哈尔滨 10 年(2002~2011 年)间的环境因子数据进行分析,其中,大豆生育期降雨量的变化范围为 356.9 mm~529.2 mm,其中 2002 年的降雨量较高,其余年份近于平稳;生育期日照时数的变化范围为 1104 h~1439.5 h,2003 年日照时数最少为 1104 h。2002 和 2004 年日照时数偏高,在 1400 h 左右,其余年份稳定在 1300 h 左右;生育期的平均气温较为平稳,在 17℃到 18℃之间,变动幅度接近于 1℃ (表 1)。

3.2. RIL 群体的表型统计分析

重组自交系群体蛋白质性状表型分析列于表 2 中。母本 Charleston 的蛋白质含量变化范围在 38.06%到 42.38%之间,父本东农 594 的蛋白质含量变化范围在 41.80%到 43.94%之间。群体的均值变化范围在 36.24%到 42.24%之间。该群体在各年份/环境的变幅较大,适合做 QE 分析。

3.3. RIL 群体蛋白性状在哈尔滨地点 10 年的品质分析

通过分析哈尔滨 10 年 10 个环境得到 GGE 双标图,其中横坐标为表型值增大的方向,纵向坐标轴

Table 1. Environment factors of Harbin from 2002 to 2011

表 1. 哈尔滨 2002~2011 年环境因子

	2002	2003	2004	2005	2006	2007	2008	2009	2010	2011
生育期降雨量 GP (mm)	529.2	466.6	447.8	408.6	454.6	356.9	368	441.3	455.5	401
生育期日照时数 ASH (h)	1377.3	1104	1439.5	1233	1347.9	1385.9	1318.2	1316.1	1308.6	1298
生育期平均气温 AT (°C)	16.9	17.4	18.1	17.5	17.8	18.1	18.2	17.4	18.3	17.9

Table 2. Summary of RIL and parents protein content

表 2. 重组自交系群体及亲本蛋白含量总结

环境	亲本 Parent			RIL 重组自交系 RIL population						
	东农 594	Charleston	均值	标准差	变异系数	偏度	峰度	最小值	最大值	变幅
Environment	dongnong594	Charleston	Average	SD	CV	Steve	Kurt	Min	Max	Range
2002HRB	42.38	40.35	39.30	1.56	0.04	0.52	0.80	35.75	44.45	8.71
2003HRB	42.95	41.21	36.24	2.52	0.07	0.04	0.52	29.86	43.55	13.69
2004HRB	43.94	42.38	36.31	1.61	0.04	0.70	1.40	31.48	42.09	10.61
2005HRB	42.68	41.21	39.66	1.48	0.04	0.74	0.38	36.84	44.71	7.87
2006HRB	41.80	40.20	39.02	1.50	0.04	0.42	-0.06	35.82	43.06	7.24
2007HRB	42.93	38.43	42.24	1.62	0.04	0.09	-0.43	38.43	46.42	7.99
2008HRB	43.04	38.06	39.17	1.57	0.04	0.42	-0.43	35.90	43.06	7.16
2009HRB	43.37	38.78	41.60	1.10	0.03	-0.02	-0.55	38.98	44.36	5.38
2010HRB	42.42	40.72	41.90	0.95	0.02	-0.24	-0.38	39.74	44.00	4.29
2011HRB	42.13	40.80	41.35	0.99	0.02	-0.09	-0.26	38.97	43.63	4.66

3.5. 最优环境因子的挖掘

如图 4 所示, SUNSHINE 的向量与目标向量夹角最小, 其对蛋白质品质影响最大; RAINFALL 的夹角略大于 SUNSHINE, 说明其对品质影响相对较小; AT 对蛋白质的影响最小。对哈尔滨 10 年的环境来说, 环境因子对品质数据的影响顺序为: 日照时数 > 降水量 > 平均温度。第 1 主成分解释了 55.8% 的效应, 第 2 主成分解释了 31% 的效应, PC1 和 PC2 共解释环境因子影响效应的 86.8%, 据此, 推断分析的可靠性较大。

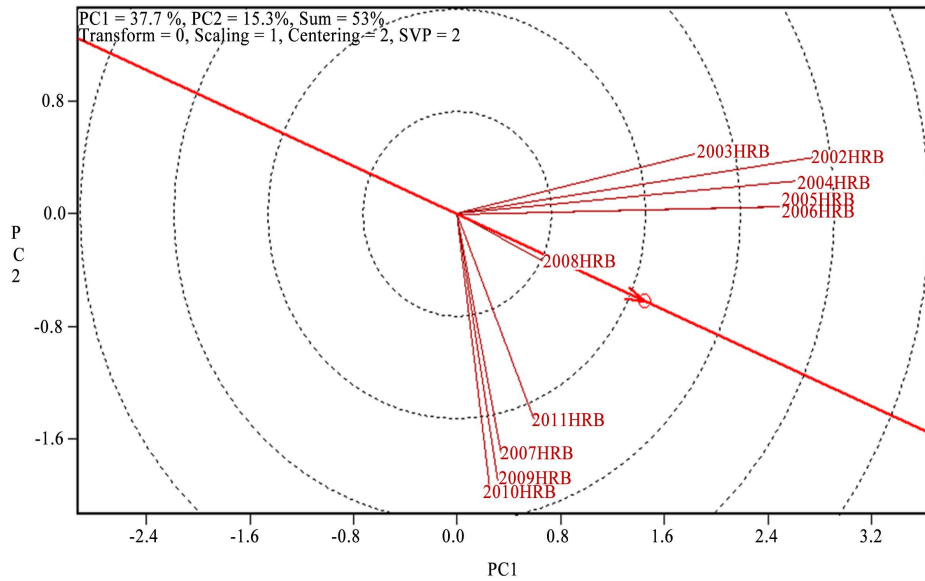
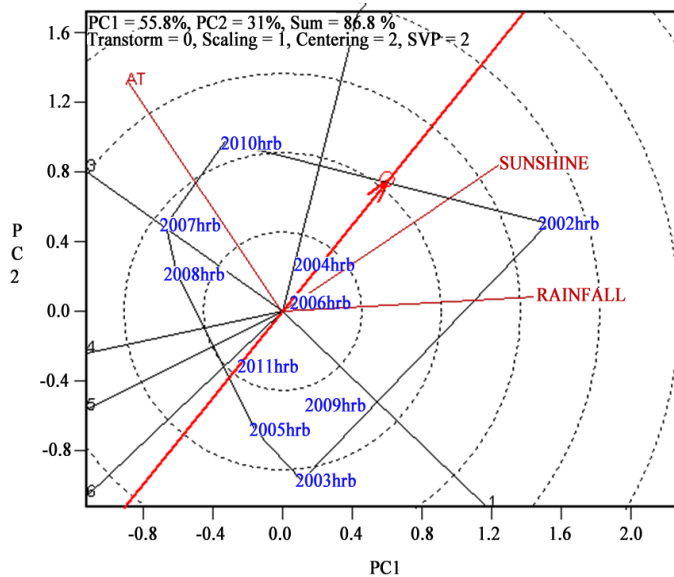


Figure 3. Identification and representativeness biplot of protein content of soybean in HRB 10 environments
图 3. RIL 群体蛋白性状在 HRB 10 环境的区分力和代表性图



注: SUNSHINE 代表大豆生育期内的日照时数, RAINFALL 代表大豆生育期内的降水量, AT 代表大豆生育期内的平均温度。

Figure 4. Environment factors GGE biplot analysis in HRB
图 4. HRB 环境的环境因子的 GGE-biplot 分析

4. 讨论

大豆蛋白质含量是数量性状, 外界环境影响因素众多, 若利用传统的二维数据, 很难解释出性状与各个环境的关系, 而 GGE 双标图法可把各环境因素与性状之间的复杂关系直观的表现出来。对多环境及多品种进行整体性评价, 能够清晰鉴别筛选出高品质性状的大豆品系, 为目标区域筛选出高产优产的大豆新品种提供参照依据, 为推广优质、优质新品种提供指导依据。

作物多点测验(或试验)是农业实验的基础, 是最经常使用的实验设计。各个省市的农作物管理部门每年都会组织各育种单位和种子公司进行各种农作物品种试验, 为育种提供依据和数据, 并通过审定。由于存在品种和环境之间的相互作用(GE), 因此多点试验是必要的。双标图法不用关注模型的适应性等问题, 可以于直接比较品种的稳定性。考虑到一年内的数据多是平衡的(相同的品种出现在所有试验点), 而年际间数据大多不平衡, 通常先每年逐一分析, 最后整合分析的思路[20]。

本研究利用 GGE 双标图法对连续 10 年的 147 个 RIL 群体的油分含量数据进行分析, 筛选出蛋白质含量较高的 5 个株系是株系 5、株系 10、株系 24、株系 80 和株系 102, 其中株系 5 和株系 10 的稳定性也较好, 这些株系的表型远优于双亲, 可以为大豆品种性状遗传育种提供丰富类型的材料, 可以进一步挖掘应用。大豆参试环境中的 2008HRB 的代表性最好, 2010HRB 代表性最差; 2002HRB 的鉴别能力最好, 2011HRB 的鉴别能力最差。环境因子对大豆蛋白质含量品质数据的影响顺序为: 日照时数 > 降水量 > 平均温度, 针对该蛋白质含量数据来说, 各环境因素中日照时数相对最重要, 应用同一群体, 通过环境的代表性和鉴别能力分析, 在连续 10 年哈尔滨环境下, 环境因子对油分含量的影响顺序为: 日照时 > 数降水量 > 平均温度, 与本文获得的研究结果完全一致[21]。另外, 王志新等[22]研究遮光对大豆种子的影响大小排序是产量 > 蛋白质 > 脂肪 > 蛋质总量, 这与本研究的日照对蛋白质的影响较大一致。GGE 双标图法能更清晰鉴别和筛选出品质性状好, 稳定性高的大豆株系, 并综合考虑基因型和基因型与环境互作的总体表现, 关注主要环境因子对大豆蛋白质含量的影响, 为指导大豆栽培生产提供科学依据和决策支持。

5. 结论

在 RIL 群体的多年环境品质性状分析中, 株系 10、23、24 与 80 的蛋白质品质好, 株系 10 具有较高的稳定性。就环境的代表性而言, 2008HRB 最好, 环境的鉴别能力是 2002HRB 最好。在哈尔滨环境下, 环境因子对品质数据的影响顺序为: 日照时数 > 降水量 > 平均温度。

参考文献

- [1] 张霞, 王峰. 植物蛋白质的特性及应用价值分析[J]. 现代农业科技, 2014(1): 289-291.
- [2] Wang, D.L., Zhu, J., Li, Z.K., *et al.* (2003) QTL Mapper Version 1.6. Copyright© by Zhejiang University, PR China, and Texas A & M University, USA.
- [3] Yang, J., Hu, C., Hu, H., *et al.* (2008) QTL Network: Mapping and Visualizing Genetic Architecture of Complex Traits in Experimental Populations. *Bioinformatics*, **24**, 721-723. <https://doi.org/10.1093/bioinformatics/btm494>
- [4] Gabriel, K.R. (1971) The Biplot Graphic Display of Matrices with Application to Principal Component Analysis. *Biometrika*, **58**, 453-467.
- [5] 严威凯, 盛庆来, 胡跃高, 等. GGE 叠图法——分析品种 × 环境互作模式的理想方法[J]. 作物学报, 2001, 27(1): 21-28.
- [6] 张勇, 何中虎, 张爱民. 应用 GGE 双标图分析我国春小麦的淀粉峰值粘度[J]. 作物学报, 2003, 29(2): 245-251.
- [7] 常磊, 柴守玺. GGE 双标图在我国旱地春小麦稳产性分析中的应用[J]. 中国生态农业学报, 2010, 18(5): 988-994.
- [8] 焦伟红, 刘景辉, 齐冰洁, 等. 用 GGE 双标图分析燕麦品种(系)农艺与品质性状[J]. 中国农学通报, 2011, 27(1): 24-29.

- [9] 周长军, 田中艳, 李建英, 等. 双标图法分析大豆多点试验中品系产量稳定性及试点代表性[J]. 大豆科学, 2011, 30(2): 318-321.
- [10] 张春明, 赵雪英. 用 GGE 双标图分析区域试验中小豆品系的高产稳产性及适应性[J]. 农学学报, 2013, 3(1): 6-9.
- [11] 梁黔云, 李清超, 吴瑞, 等. GGE 双标图在玉米品种区域实验中的应用[J]. 中国农学通报, 2014, 30(6): 224-228.
- [12] 许乃银, 张国伟, 李健, 等. 基于 HA-GGE 双标图的长江流域棉花区域试验环境评价[J]. 作物学报, 2012, 38(12): 2229-2236.
- [13] 罗俊, 张华, 邓祖湖, 等. 应用 GGE 双标图分析甘蔗品种(系)的产量和品质性状[J]. 作物学报, 2013, 39(1): 142-152.
- [14] 张勇, 何中虎. 应用 GGE 双标图分析我国春小麦的淀粉峰值粘度[J]. 作物学报, 2003, 29(2): 245-251.
- [15] 罗俊, 张华, 邓祖湖, 等. 用 GGE 双标图分析甘蔗品种性状稳定性及试点代表性[J]. 应用生态学报, 2012, 23(5): 1319-1325.
- [16] 杨进文, 朱俊刚, 王曙光, 等. 用 GGE 双标图及隶属函数综合分析山西小麦地方品种抗旱性[J]. 应用生态学报, 2013, 24(4): 1031-1038.
- [17] 张春明, 赵雪英. 用 GGE 双标图分析区域试验中小豆品系的高产稳产性及适应性[J]. 农学学报, 2013, 3(1): 6-9.
- [18] 韩英鹏, 孙明明, 刘昊飞, 等. GGE 双标图法评价大豆重组自交系亚麻酸表现的稳定性[J]. 大豆科学, 2014(5): 776-778.
- [19] 韩英鹏, 赵雪, 滕卫丽, 等. GGE 双图标法评价大豆脂肪酸表现的稳定性[J]. 大豆科学, 2014(4): 514-518.
- [20] Song, X., Han, Y., Teng, W., *et al.* (2010) Identification of QTL Underlying Somatic Embryogenesis Capacity of Immature Embryos in Soybean (*Glycine max* (L.) Merr.). *Plant Cell Reports*, **29**, 125-131.
<https://doi.org/10.1007/s00299-009-0804-1>
- [21] 齐慧冬, 齐照明, 侯萌, 等. 基于 GGE 双标图的大豆油分多环境的稳定性分析[J]. 大豆科学, 2016, 35(1): 11-17.
- [22] 王志新. 环境因素对大豆化学品质及产量影响研究 II. 遮光对大豆化学品质及产量的影响[J]. 大豆科学, 2004, 23(1): 41-44.

知网检索的两种方式:

1. 打开知网页面 <http://kns.cnki.net/kns/brief/result.aspx?dbPrefix=WWJD>
下拉列表框选择: [ISSN], 输入期刊 ISSN: 2164-5507, 即可查询
2. 打开知网首页 <http://cnki.net/>
左侧“国际文献总库”进入, 输入文章标题, 即可查询

投稿请点击: <http://www.hanspub.org/Submission.aspx>

期刊邮箱: hjas@hanspub.org