

小麦品种生态类型及其演变的统计分析方法研究*

袁志发

宋哲民

刘光祖

(西北农学院基础课部) (西北农学院农学系) (西北农学院基础课部)

摘 要

作物品种生态型的基本特点及其演变规律,对制定作物育种目标、亲本选配、杂种后代选择,以及品种适应性鉴定等一系列基本育种环节,都具有极为重要的指导意义。因此品种生态类型的研究,日益引起育种工作者的关注。

本文在前人研究的基础上,给出了生态型基本特征和可变特征的含义,并给出了区分它们的初步的统计方法——单变量分析法和多变量分析法。

单变量分析法运用方差分析和多重比较来判断品种性状的区分能力,以确定性状的稳定性,运用Nair水平分组法给出各性状的定量的区分类型。

多变量分析法运用R型因子分析来判断品种生态型的基本特征和可变特征。相关矩阵的第一主成分代表品种生态型的基本特征,第二、第三等主成份代表生态型的可变特征。按主因子进行聚类来区分生态型的亚型。运用Q型因子分析判断生态地区的代表品种。

多变量分析方法比单变量分析方法获得的信息丰富、集中,且便于分类和解释。

※

※

※

一、前 言

作物品种生态类型的基本特点及其演变规律,对制定作物育种目标、亲本选配、杂种后代选择,以及品种适应性鉴定等一系列基本育种环节,都具有极为重要的指导意义。在一定意义上,品种适应性的好坏,育种工作成效的大小,很大程度取决于对当地品种生态类型及其演变规律认识的深度和广度。因此,品种生态类型的研究,日益引起育种工作者的关注。

陕西关中地区的小麦品种生态类型,王玉成、赵洪璋(1947)曾划分为五个生态亚型^[1];金善宝等(1959、1961)在全国生态类型的研究中,划入华北平原生态类型,并分为三个亚型^[2]。以后,在有关全国小麦品种生态分类中,根据新的情况和资料,将关中地区的小麦品种称为黄淮平原生态类型^[3]。育种实践表明,这些研究工作对关

*本刊编辑室收到此稿时间:1984年6月30日。

中地区小麦育种, 已经并将继续起重要的指导作用。

关于品种生态类型的基本特点及演变问题, 近几年, 赵洪璋^[4]、李振声^[5]等许多研究者从不同角度进行了探讨, 提出了一些值得深入研究的问题。综析前人的研究, 我们认为: 对一个地区小麦品种生态型及其演变的研究, 可建立如下的概念和相应的研究方法。

1. 生态型的基本特征。生态型的基本特征是指作物品种受本地区生态条件长期制约、自然的选择和人工选择形成的基本生态特征, 具有相对稳定性, 如冬性、春性等。一个地区的生态型的基本特征中, 除了明显的如冬春性之类的性状外, 还有哪些性状? 这个问题的解决显然是必要的。

2. 生态型的可变特征。随着本地区重要栽培条件的改变, 小麦品种的基本特征和其它特征将会向适应于新的栽培条件的变化而发生不同程度的变化, 这种变化将受基本特征稳定性的约束、协调达到新的统一。显然这些信息的获得对育种工作有现实意义。

本文在于探索区分生态型的基本特征及可变特征的统计分析的初步方法。

二、试验材料与试验设计

设当地的农家品种有 K_0 个: $A_{01}, A_{02}, \dots, A_{0K_0}$ 。本地区栽培条件的重大改变有 l 次: $t_1, t_2, \dots, t_l$ 。在每个生态阶段 t_i 中所出现的新品种有 K_i 个: $A_{i1}, A_{i2}, \dots, A_{iK_i}$ ($i=1, 2, \dots, l$)。目前待鉴定的品种有 K_{l+1} 个: $A_{l+1,1}, A_{l+1,2}, \dots, A_{l+1,K_{l+1}}$ 。共有材料 $\sum_{i=0}^{l+1} K_i = K$ 个。每一品种用编号 j 表示 ($j=1, 2, \dots, K$)。按随机区组进行试验, r 次重复。各品种为单行区, 每行15株左右。测定的性状可根据生长阶段按所需酌定。

用 X_{pj} 表示第 j 个品种第 p 次重复的某性状的小区平均值, 则试验资料符号如表1所示。

表1 资 料 符 号 表

重 复	品 种						
		1	2	...	j	...	K
1		X_{11}	X_{12}	...	X_{1j}	...	X_{1k}
2		X_{21}	X_{22}	...	X_{2j}	...	X_{2k}
$\vdots$		$\vdots$	$\vdots$	...		...	$\vdots$
P		X_{p1}	X_{p2}	...	X_{pj}	...	X_{pk}
$\vdots$		$\vdots$	$\vdots$	...		...	$\vdots$
r		X_{r1}	X_{r2}	...	X_{rj}	...	X_{rk}

三、单变量统计分析方法

由生态型的基本特征的含意可知,在统计上某性状对品种无区别能力或区别能力甚差,则此性状为生态型的基本特征,否则应视为可变特征。为此,对每一性状进行方差分析,即:

1.进行F检验,若不显著,则此性状为基本特征。

2.若F检验显著,应用q法进行多重比较,再按奈尔(Nair)方法^[6]进行水平分组,以确定该性状的分组和区分品种的能力。

下面以西北农学院小麦育种研究室1981年品种区域试验总结中的资料为例,说明这种分析方法。

经对原资料整理得早肥组小麦各性状的方差分析(表2)。

表2 早肥组品种各性状的方差分析

品 种 号	性状	冬季分蘖	株高	每穗粒数	千粒重	抽穗期	成熟期
	均值						
1	(小偃6号)	11.5	95.3	26.4	39.2	4/19	6/2
2	(7576/3矮7)	9.0	97.7	30.8	46.8	4/17	6/6
3	(68G(2)8)	7.9	110.7	39.7	39.1	4/17	6/3
4	(7019 0-1)	9.1	89.0	35.4	35.3	4/18	6/2
5	(9615-11)	11.6	88.0	29.3	37.0	4/20	6/7
6	(9615-13)	13.0	87.7	24.6	44.8	4/19	6/7
7	(73 (36))	11.6	79.7	25.6	43.7	4/19	6/4
8	(丰3)	10.7	119.3	29.9	38.8	4/19	6/5
9	(矮3)	11.1	87.7	32.2	35.6	4/18	6/3
机误均方 S^2_e		1.7	8.9	5.6	0.72	0.15	0.6
F 值		4.6**	53.6**	12.6**	73.3**	21**	8.33**

表2结果说明各性状的F检验都是极显著的,用奈尔方法进行水平分组来确定各性状的分组。奈尔方法的步骤用冬季分蘖说明之:

①选取显著性水平 α ,我们取 $\alpha = 0.05$

②计算

$$d_i = t_{\alpha}(f_e) \sqrt{\frac{2S^2_e}{r}}$$

这里 f_e 为机误均方 S^2_e 的自由度, r 为重复数, t_{α} 是自由度为 f_e 的 t 临界值。这里 $r = 3$,

$fe=16, \quad t_{0.05}=2.120, \quad S^2_e=1.7, \quad d_1=2.120\sqrt{\frac{2 \times 1.7}{3}}=2.3$

③把各品种均值按大小次序排队，并计算相邻二均值差，然后与 d_1 比较，凡比 d_1 大的分为一组。比 d_1 小的分为一组。关于冬季分蘖的初步分组如表3，因相邻差数均小于 d_1 ，故只能分为一组。

表3 旱肥组品种冬季分蘖的奈尔分组法

分组结果	一 组								
品种号	6 (9615—13)	5 (9615—11)	7 (73<36>)	1 (小偃6号)	9 (矮丰3号)	8 (丰产3号)	4 (7019-0-1)	2 (7576/3矮7)	8 (68G<2>8)
均值	13	11.6	11.6	11.5	11.1	10.7	9.1	9.0	7.9
相邻差数	1.4	0	0.1	0.4	0.4	1.6	0.1	1.1	

④如果按上面的分组，每一组都不超过两个品种，则分组工作就到此结束；如果有的组超过两个品种，对这些组再进一步考查。上面只分了一个组，包括9个品种，因而需再考查。

⑤对超过两个品种的组进行考查。令 q 表示该组的品种数， $\bar{X}_q$ 表示这 q 个均值的平均， dq 表示 q 个均值与 $\bar{X}_q$ 之差的绝对值最大者，再计算

$$U = \begin{cases} \frac{dq\sqrt{\frac{r}{S^2_e}} - \frac{1}{2}}{3\left(\frac{1}{4} + \frac{1}{f_e}\right)}, & \text{当 } q=3, \\ \frac{dq\sqrt{\frac{r}{S^2_e}} - \frac{6}{5}\lg q}{3\left(\frac{1}{4} + \frac{1}{f_e}\right)}, & \text{当 } q>3 \end{cases}$$

在 $\alpha=0.05$ 之下，当 $U>1.96$ 时，则将 该组离 $\bar{X}_q$ 最远的那个均值剔出这个组，如此逐个剔除，直到 $U<1.96$ 为止。对于剔除掉的品种另成一组，再按此法确定是否剔除，是否需另外成立新组，直到各组无剔除的品种为止，分组 就告 完成。对于上例， $q=9$ ，

$\bar{X}_q = \bar{X}_9 = \frac{1}{9} (13+11.6+\cdots+7.9) = 10.6$

$d_9 = d_9 = 2.7$

$$U = \frac{2.7\sqrt{\frac{3}{1.7}} + \frac{6}{5}\lg 9}{3\left(\frac{1}{4} + \frac{1}{16}\right)} = \frac{3.5867 - 1.1451}{0.9375} = 2.604 > 1.96$$

由于 $10.6 - 7.9 = 2.7$, $13 - 10.6 = 2.4$, 故应剔除品种3。因此分组情况变为二组:

13 11.6 11.6 11.5 11.1 10.7 9.1 9.0
7.9

又第一组 $q = 8$, $\bar{X}_8 = 10.95$, $d_8 = 13 - 10.95 = 2.05$, 而

$$U = \frac{2.05\sqrt{\frac{3}{1.7}} - \frac{6}{5}\lg 8}{3\left(\frac{1}{4} + \frac{1}{16}\right)} = \frac{2.7233 - 1.0837}{0.9375} = 1.7489 < 1.96$$

这说明上述分组从统计观点已经确定。如果将品种6除, 则分组亦满足统计学要求。这时分组为:

13
11.6 11.6 11.5 11.1 10.7 9.1 9.0, 7.9 ($\bar{X}_8 = 10.3$)

从生产实践来看第二种分组是合适的。

多重比较的方法可参考有关书籍^[6]。我们用 q 法, 取 $\alpha = 0.01$, 则冬季分蘖的多重比较结果为

品种号:	6	5	7	1	9	8	4	2	3
均值:	13	11.6	11.6	11.5	11.1	10.7	9.1	9.0	7.9
比较结果:	A	AB	AB	AB	AB	AB	AB	AB	B

一般来讲, 如果参加试验的品种为 K 个, 则需相互比较 $\frac{K(K-1)}{2}$ 对。这里 $K = 9$,

应比较36对。从比较结果看, 只有一对(6号与3号)是显著的, 因此冬季分蘖区分品种的能力为 $1/36$ 。

上述试验(早肥组)各性状的奈尔分组结果及区分品种能力如表4所示。

表4 早肥组品种各性状奈尔分组及区别品种能力

冬 季 分 蘖	品种号	6 (9615 -13)	5 (9615 -11)	7 (73 (36))	1 (小偃 6号)	9 (矮丰 3号)	8 (丰产 3号)	4 (7019 -0-1)	2 7576/ 3矮7	3 (68G (2)8)
	均 值	13	11.6	11.6	11.5	11.1	10.7	9.1	9.0	7.9
	分 组	多蘖	中 少 蘖							少蘖
	组均值	13	10.7							7.9
	区分品 种能力	1/36 = 0.028								

续表4

株高	品种号	8	3	2	1	4	5	6	9	7
	均值	119.3	110.7	97.7	95.3	89	88	87.7	87.7	79.9
	分组	高秆	中秆	中矮秆		矮秆				
	组均值	119.3	110.7	96.5		86.4				
	区分品种能力	23/30 = 0.61								
每穗粒数	品种号	3	4	9	2	8	5	1	7	6
	均值	39.7	35.4	32.2	30.8	29.9	29.3	26.4	25.6	24.6
	分组	多粒	中多粒	少粒						
	组均值	39.7	35.4	29.3						
	区分品种能力	10/36 = 0.28								
千粒重	品种号	2	6	7	1	3	8	5	9	4
	均值	46.8	44.8	43.7	39.2	39.1	38.8	37	35.6	35.3
	分组	重粒	大粒		中粒			小粒		
	组均值	46.8	44.8		39.0			33.0		
	区分品种能力	25/36 = 0.69								
抽穗期	品种号	5	6	7	1	8	4	9	2	3
	均值	4/20	4/19	4/19	4/19	4/19	4/18	4/18	14/17	14/17
	分组	晚	中晚				中		早	
	组均值	20	19				18		17	
	区分品种能力	28/36 = 0.78								
成熟期	品种号	6	5	2	8	7	3	9	1	4
	均值	6/7	6/7	6/6	6/5	6/4	6/3	6/3	6/2	6/2
	分组	晚		中		早				
	组均值	7		5.5		2.8				
	区分品种能力	14/36 = 0.39								

关于水肥组试验各性状的方差分析结果如表5所示。奈尔分组结果如表6所示。

表5 水肥组品种各性状的方差分析

品种号 \ 性状 均值	冬季分蘖	株高	每穗粒数	千重粒	抽穗期	成熟期
1 (7410—RO)	11.7	100.3	25.8	41.4	4/21	6/5
2 (7576/3矮7)	10.8	100.7	27.6	49.5	4/18	6/7
3 (73(15) - 5 - 1)	7.3	85.0	38.6	40.5	4/10	6/4
4 (9615) - 11)	11.2	91.3	28.2	38.0	4/21	6/8
5 (9617 - 13)	13.0	91.3	22.3	45.2	4/20	6/7
6 (偃师4号)	10.1	88.3	30.4	46.4	4/11	6/2
7 (73<36>)	9.9	82.0	25.2	47.2	4/21	6/7
8 (77(2))	8.0	91.7	29.6	44.4	4/13	6/5
9 (阿勃)	8.6	120.3	30.6	39.6	4/20	6/7
10 (郑引1号)	8.5	102.0	38.9	37.2	4/16	6/5
11 (矮丰3号)	10.4	87.3	30.7	36.8	4/19	6/6
机误均方	0.98	3.78	4.8	1.1	0.15	0.60
F 值	8.92**	92.0**	16.4**	52.1**	21.0**	8.33**

表6 水肥组品种各性状的奈尔分组及区分品种的能力

冬季分蘖	品种号	5	1	4	2	11	6	7	9	10	8	3
	均值	13.0	11.7	11.2	10.8	10.4	10.1	9.9	8.6	8.5	8.0	7.3
	分组	多蘖	中蘖	中少蘖							少蘖	
	组均值	13.0	11.7	9.9							7.7	
	区分品种能力	6/55 = 0.11										

株高	品种号	9	10	2	1	8	4	5	6	11	3	7
	均值	120.3	102	100.7	100.3	91.7	91.3	91.3	88.6	87.3	85.0	82.0
	分组	高秆	中秆			中矮秆					矮秆	
	组均值	120.3	101.0			90.0					83.5	
	区分品种能力	37/55 = 0.67										

续表6

每穗粒数	品种号	10	3	11	9	6	8	4	2	1	7	5
	均值	33.9	33.5	30.7	30.6	30.4	29.6	28.2	27.6	25.8	25.2	22.3
	分组	多粒		中粒								少粒
	组均值	38.75		28.5								22.3
	区分品种能力	22/55 = 0.40										
千粒重	品种号	2	7	6	5	8	1	3	9	4	10	11
	均值	49.5	47.2	46.4	45.2	44.4	41.4	40.5	39.6	38.0	37.2	36.8
	分组	重粒	大粒				中粒			小粒		
	组均值	49.5	45.8				40.5			37.3		
	区分品种能力	36/55 = 0.65										
抽穗期	品种号	1	4	7	5	9	11	2	10	8	6	3
	均值	4/21	4/21	4/21	4/20	4/20	4/19	4/18	4/16	4/13	4/11	4/10
	分组	按天分组										
	组均值											
	区分品种能力	41/55 = 0.75										
成熟期	品种号	4	2	5	7	9	11	8	10	1	3	6
	均值	6/8	6/7	6/7	6/7	6/7	6/6	6/5	6/5	6/5	6/4	6/2
	分组	晚					中				早	
	组均值	7.2					5.3				3	
	区分品种能力	32/55 = 0.58										

据上述分析,由各性状区分品种的能力可看出它们的稳定性,这对我们确定生态型的基本特征和可变特征提供了依据。各性状稳定性的位次如表7所示。

表7 各性状稳定性的位次

试验型	性状	冬季分蘖	每穗粒数	成熟期	株高	千粒重	抽穗期
	位次						
早肥组	区分品种能力	0.028	0.28	0.39	0.64	0.69	0.78
	位次	1	2	3	4	5	6
水肥组	区分品种能力	0.11	0.40	0.58	0.67	0.65	0.75
	位次	1	2	3	5	4	6

表8 各品种的生态学特征

品 种 \ 特 征	冬季分蘖	株 高	成熟期	每穗粒数	千粒重	抽穗期	冬春性
7014-Ro	中-中少	中矮	中早	少	中	中晚	弱冬
7576/3矮7	中少	中	中晚	中少	重	早	半冬
68G<2>8	少	中	早	多	中	早	弱冬
7319-0-1	中少	矮	早	中多	小	中	弱冬
9615-11	中少	矮	晚	中少	小	晚	冬
9617-13	多	矮	晚	少	大	中晚	冬
73<36>	中少	矮	中	少	大	中晚	弱冬
丰产3号	中少	高	中	少	中	中晚	弱冬
矮丰3号	中少	矮	中早	中多	小	中	弱冬
73<15>5-1	少	矮	早	多	中	早	半春
偃师4号	中少	矮	早	中多	大	早	春
77<2>	少	中矮	中	中多	大	早	半春
阿勃	中少	高	晚	中多	中	中晚	半春
郑引1号	中少	中	中	多	小	中	半春

如果说参加试验的品种能代表陕西关中地区小麦的生态类型,则上述分析中可以给人们如下的初步印象:从品种的异质性来讲,主要起作用的是株高、千粒重和抽穗期,而

前两个位次的性状主要说明品种的同源性。从生态学观点看,前面位次的性状,尤其是分蘖特性是品种生态型的基本特征,而后三个性状是可变的性状或者说它们的可塑性较大。也可以说,品种在分蘖、每穗粒数、成熟期等方面没有什么重大突破,而其它三个性状变化很大。

下面我们对上述试验中各个品种作一个初步的生态学的描述(如表8所示)。

四、多变量分析方法

上述单因素分析方法虽简单,但有一些缺点,如它未考虑品种各性状之间或强或弱的相关性,因此,须进行多变量分析。多变量分析方法回答以下几个问题:

①在本地区内推广的小麦品种中,起支配作用的性状是什么?即要回答生态型的基本特征和可变特征。

②本地区小麦品种的生态型中,存在着那些小的类别?

③本地区小麦品种中的代表品种是什么?

为了回答此问题,可采用R型与Q型因子分析方法。

R型因子分析是研究小麦品种各性状间的相关关系的,通过相关矩阵内部结构的研究,找出控制所有性状的几个主成分,这几个主成分是由各性状的不同线性组合成的综合性状,它们之间是相互独立的。每个主成分不同程度的代表着支配所有品种的几个性状或一个性状,它反映了小麦品种各性状间结合的一种遗传的和生态的作用。R型因子分析可以帮助我们回答上述的前两个问题。

Q型因子分析与R型因子分析的方法是一样的,不过Q型因子分析是研究各品种间相关关系的,可以找出各品种中共同的代表性品种和脱离代表性品种的两极品种。

关于旱肥组的因子分析

1. R型因子分析

①计算各性状间的相关系数,组成相关矩阵:

	分蘖	株高	粒数	粒重	抽穗期	成熟期
分蘖	1	-0.4818	-0.8875	0.1456	0.8123	0.4044
株高		1	0.4369	-0.0853	-0.2979	-0.0508
粒数			1	-0.4709	-0.6883	-0.4320
粒重				1	-0.1653	0.5148
抽穗期					1	0.3493
成熟期						1

②计算R的特征根 $\lambda_1 \geq \lambda_2 \geq \lambda_3 \geq \lambda_4 \geq \lambda_5 \geq \lambda_6$, 求出相应于各特征根的特征向量, 并组成特征向量矩阵 $U = (U_1, U_2, U_3, U_4, U_5, U_6)$ 。

$$\lambda = \begin{matrix} 3.16563 & 1.38500 & 0.85953 & 0.45758 & 0.107731 & 0.02446 \end{matrix}$$

$$U = \begin{pmatrix} 0.52684 & -0.18910 & 0.05834 & -0.16538 & -0.69630 & 0.41363 \\ -0.30111 & 0.23946 & 0.78550 & -0.47171 & -0.11152 & 5.88 \times 10^{-3} \\ -0.52947 & -0.05380 & 0.06802 & 0.44782 & -0.06344 & 0.71244 \\ 0.20741 & 0.73484 & -0.28027 & -0.28220 & 0.26048 & 0.43697 \\ 0.44506 & -0.37397 & 0.39169 & 0.02971 & 0.64455 & 0.30384 \\ 0.33217 & 0.47345 & 0.37814 & 0.68489 & -0.12410 & -0.19504 \end{pmatrix}$$

③计算因子载荷矩阵 $A = [\sqrt{\lambda_1} U_1, \sqrt{\lambda_2} U_2, \sqrt{\lambda_3} U_3, \sqrt{\lambda_4} U_4, \sqrt{\lambda_5} U_5, \sqrt{\lambda_6} U_6] = (a_{ij})$ 。

A的元素 a_{ij} 叫做性状 X_i 在 F_j 上的载荷，A可表示如下：

$$A = \begin{matrix} & \begin{matrix} F_1 & F_2 & F_3 & F_4 & F_5 & F_6 \end{matrix} \\ \begin{matrix} X_1 \\ X_2 \\ X_3 \\ X_4 \\ X_5 \\ X_6 \end{matrix} & \begin{pmatrix} a_{11} & a_{12} & a_{13} & a_{14} & a_{15} & a_{16} \\ a_{21} & a_{22} & a_{23} & a_{24} & a_{25} & a_{26} \\ a_{31} & a_{32} & a_{33} & a_{34} & a_{35} & a_{36} \\ a_{41} & a_{42} & a_{43} & a_{44} & a_{45} & a_{46} \\ a_{51} & a_{52} & a_{53} & a_{54} & a_{55} & a_{56} \\ a_{61} & a_{62} & a_{63} & a_{64} & a_{65} & a_{66} \end{pmatrix} \end{matrix}$$

$$F_j = a_{1j}X_1 + a_{2j}X_2 + a_{3j}X_3 + a_{4j}X_4 + a_{5j}X_5 + a_{6j}X_6, \quad j = 1, 2, \dots, 6$$

$$X_i = a_{i1}F_1 + a_{i2}F_2 + a_{i3}F_3 + a_{i4}F_4 + a_{i5}F_5 + a_{i6}F_6, \quad i = 1, 2, \dots, 6$$

各主因子 F_j 的方差贡献为 λ_j ， λ_j 愈大， F_j 所反映的性状组合所代表生态学作用愈稳定。具体计算结果为：

$$\begin{matrix} & F_1 & F_2 & F_3 & F_4 & F_5 & F_6 \\ \begin{matrix} \text{分蘖}-X_1 \\ \text{株高}-X_2 \\ \text{粒数}-X_3 \\ \text{粒重}-X_4 \\ \text{抽穗期}-X_5 \\ \text{成熟期}-X_6 \end{matrix} & \begin{pmatrix} 0.93736 & -0.22254 & 0.05408 & -0.11190 & -0.22854 & 0.06441 \\ -0.53575 & 0.28181 & 0.72824 & -0.31918 & -0.03660 & 0.00092 \\ -0.94205 & -0.06332 & 0.06306 & 0.30302 & -0.02082 & 0.11093 \\ 0.36902 & 0.86482 & -0.25984 & -0.19095 & 0.08549 & 0.06804 \\ 0.79185 & -0.44015 & 0.36314 & 0.02010 & 0.21155 & 0.04731 \\ 0.59101 & 0.55718 & 0.35058 & 0.46342 & -0.04073 & -0.03037 \end{pmatrix} \end{matrix} = A$$

方差贡献累计(%)：53 75.8 90.2 97.8 99.6 100

由上述的结果可看出，前三个主因子是重要的。因为它们的方差贡献为90.2%，其生物学意义如下：

F_1 ：反映了各性状组合成品种的情况，反映了品种生态型内性状组合的稳定性，由 F_1 看出，参试品种基本上以分蘖×粒数×抽穗期×成熟期为主。

F_2 ：基本组分为粒重，说明它是可变的。

F_3 ：主要组分为株高，说明它是可变的。

因此，第一主因子 F_1 反映了性状组合的稳定性，即分蘖、粒数、抽穗期、成熟期为生态

型的基本特征, 另外, 粒重和株高分别是 F_2 、 F_3 的主要组分, 它们的变化形成了各品种的主要差异, 反映了生态型的可变特性。以上分析和单因素分析的结果基本一致。

④计算因子得分。为了对品种进行分类和进一步的解释, 可利用 $F_j = \sum_{i=1}^6 a_{ij} X_i$

(由表2数据) 计算各品种的主因子得分, 其结果如表9所示。

表9 旱肥组各品种的因子得分表

	F_1	F_2	F_3	F_4	F_5	F_6
1	10.77964	-2.55921	0.62192	-1.28685	-2.62821	0.74072
2	8.43624	-2.00286	0.48672	-1.0071	-2.05686	0.57969
3	7.40514	-1.75807	0.42723	-0.88401	-1.80547	0.50884
4	8.529976	-2.02511	0.49213	-1.01829	-2.07977	0.58613
5	10.87337	-2.58146	0.62733	-1.29804	-2.65106	0.74716
6	12.18569	-2.89302	0.70304	-1.45470	-2.97102	0.83733
7	10.87337	-2.58146	0.62733	-1.29804	-2.65106	0.74716
8	10.02975	-2.38118	0.57866	-1.19733	-2.44538	0.68919
9	10.40470	-2.47019	0.60029	-1.24209	-2.53679	0.71495

从表9中可以看出: 各品种在6个主因子得分的趋势相当一致, 这说明参试各品种基本上属于一个大的生态型, 它们之间仅大同小异而已。另外品种5与品种7相应的得分完全一样, 这说明克服了相关性以后, 品种间的相似性更清楚了。为了直观起见, 取 F_1 、 F_2 两个轴作因子得分(图1)。

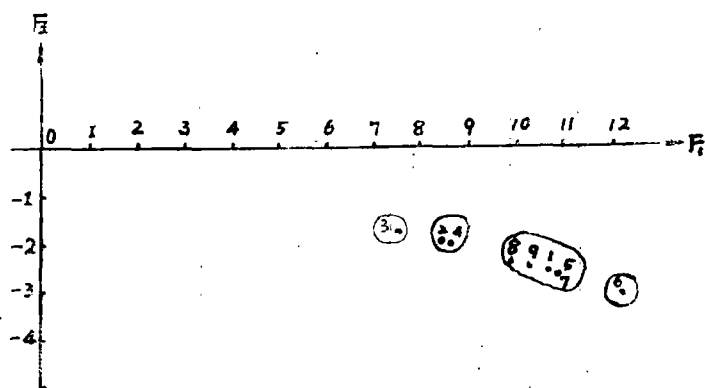


图1 旱肥组因子得分图

由因子得分图可看出, 9个品种基本属于一个类型, 其特征 F_1 轴得正分, F_2 轴得负分。如果再细分, 直观上可分为4类。

为了从数量上进行聚类分析，可按下列公式计算第*i*、*j*两个品种之间的距离

$$D_{ij} = \sum_{k=1}^6 (F_{ki} - F_{kj})^2$$

D_{ij}^2 可解释为品种间生态和遗传上的距离。

我们现取*F*₁、*F*₂、*F*₃计算各品种间的距离，其结果如表10所示。

表10 D^2 表

品种	3	2	4	8	9	1	5	7	6
3	0	1.12661	1.3408	7.2997	9.5344	12.0670	12.7466	12.7466	24.2178
2		0	0.0093	2.6909	4.1061	5.8193	6.2942	6.2942	14.4975
4			0	2.9836	3.7244	5.3631	5.8193	5.8193	14.1619
8				0	0.1490	0.5959	0.7542	0.7542	4.9263
9					0	0.1490	0.2328	0.2328	3.3613
1						0	0.0093	0.0093	2.0950
5							0	0	1.8249
7								0	1.8249
6									0

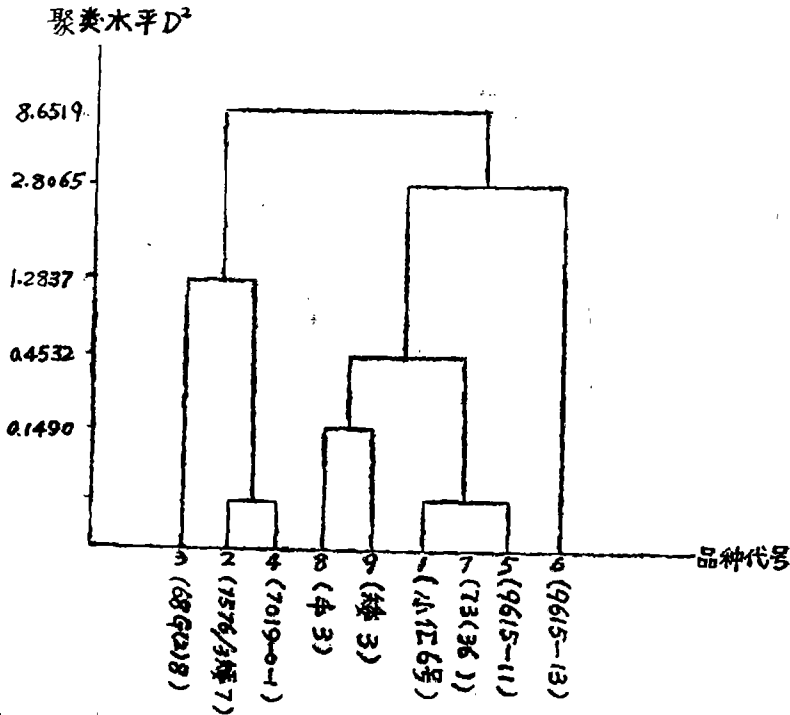


图2 品种聚类图

由表10可进行聚类分析。开始每个品种自成一类，每两品种间的距离，即为开始间的距离。然后把距离最小的两类并为一个新类，新类到其它各类的距离，采用“类平均法”来计算。以上聚类过程，每聚一次，类数减少一个，直到所有品种并成一类为止。根据聚类过程，按照聚类的次序和水平可画出品种聚类图（如图2）。

当聚类水平 $D^2 = 0.4532$ 时分成四类：

第Ⅰ类：{3}类内距离为0

第Ⅱ类：{2, 4}类内距离为0.0093

第Ⅲ类：{8, 9, 1, 7, 5}类内距离为0.2887

第Ⅳ类：{6}类内距离为0

各类间距离分别为：

$$D^2_{I \ II} = 1.2837$$

$$D^2_{I \ IV} = 14.3297$$

$$D^2_{I \ III} = 10.8787$$

$$D^2_{II \ IV} = 2.8065$$

$$D^2_{I \ IV} = 24.278$$

$$D^2_{II \ III} = 4.8314$$

这里类内距离都小于类间距离，各类呈自然聚集状态，分类是较合理的。

为了更清楚地看出各类品种的性状组合规律，结合单因子分析结果，整理得各类别性状组合规律（表11）。

表11

旱肥组各类品种性状的组合规律

类别	性状						
		冬季分蘖	株高	每穗粒数	千粒重	抽穗期 (4月)	成熟期 (6月)
I	3	少	中	多	中	17	3
II	4	中少	矮	中多	小	18	2
	2	中少	中矮	中少	重	17	6
III	9	中	矮	中	小	18	3
	7	中多	矮	少	大	19	4
	1	中多	中矮	少	中	19	2
	8	中少	高	中少	中	19	5
	5	中多	矮	少	小	20	7
IV	6	多	矮	少	大	19	7
按类序性状动态		↓ 少 ↓ 多		↑ 多 ↑ 少		↓ 早 ↓ 晚	↓ 早 ↓ 晚

表11中所列4类, 大致的趋势是: 在分蘖上是由少到多的, 在每穗粒数上是由多到少, 在抽穗期与成熟期上是由早到晚的, 而其它性状是参差不齐的。

2. Q型因子分析

9个品种的相关系数很接近, 在0.9584与0.9985之间, 其第一个特征根为 $\lambda_1 = 8.924$, 占方差贡献的99%, 说明这些品种基本上属于一个生态型。相应的特征向量为

$U'_1 = (0.33406, 0.33245, 0.33204, 0.33258, 0.33462, 0.33299, 0.33310, 0.33392, 0.33423)$ 。9个品种按分量大小排队, 次序为:

分量	0.33462	0.33423	0.33406	0.33392	0.33310	0.33299	0.33258	0.33245	0.33204
品种号	5	9	1	8	7	6	4	2	3

结合得分图看出, 两极品种为5和3: 5代表中少分蘖, 粒数少、粒重小、抽穗晚品种; 而品种3代表分蘖最少, 粒数最多、粒重中等、早抽穗、早成熟品种。居中品种为1、8即7014-RO和丰产3号, 可视为代表品种。

关于水肥组因子分析结果

1. R型因子分析

①相关矩阵:

$$\begin{pmatrix} 1 & -0.091514 & -0.79569 & 0.23394 & 0.63756 & 0.38001 \\ & 1 & 0.08229 & -0.21209 & 0.28476 & 0.23200 \\ & & 1 & -0.52134 & 0.63370 & -0.46670 \\ & & & 1 & -0.10984 & -0.06517 \\ & & & & 1 & 0.81592 \\ & & & & & 1 \end{pmatrix}$$

②特征根与特征向量矩阵:

$$U = \begin{pmatrix} \lambda = 2.92859433 & 1.6039984 & 0.776132518 & 0.490752556 \\ 0.487483228 & -0.193867348 & -0.188279678 & -0.596593831 \\ 0.0691489336 & 0.550518575 & 0.764793241 & -0.300966171 \\ -0.51658617 & 0.296849624 & -0.106672826 & 0.0819360362 \\ 0.146470498 & -0.613761123 & 0.588111167 & 0.322229478 \\ 0.520952914 & 0.29634073 & -0.118886543 & 0.0208530982 \\ 0.444827731 & 0.326655332 & -0.0908527068 & 0.66521798 \end{pmatrix}$$

0.124738676	0.0757835291
0.577258283	-0.0132284325
0.0623010819	-0.113035221
0.59772888	0.519275082
0.202661308	0.333264885
-0.341273832	0.713969709
0.384801509	-0.310778387

③因子载荷矩阵:

	F ₁	F ₂	F ₃	F ₄	F ₅	F ₆	
分蘖—X ₁	0.83424	-0.24553	-0.16537	-0.41793	0.20333	-0.00364	} = A
株高—X ₂	0.11834	0.69636	0.67376	-0.21084	0.02200	-0.03112	
粒数—X ₃	-0.88405	0.37549	0.09397	0.05740	0.21111	0.14295	
粒重—X ₄	0.25066	-0.77635	0.51811	0.22574	0.07158	0.00174	
抽穗期—X ₅	0.89151	0.37481	0.10474	0.01461	0.12053	0.19655	
成熟期—X ₆	0.76124	0.41319	-0.03004	0.46601	0.13591	-0.02555	
方差贡献累计(%)	48.8	75.5	88.5	96.7	98.7	100	

F₁——反映了各性状如何组成品种的情况,基本上以分蘖×粒数×抽穗期×成熟期为主,反映了生态型内的稳定性。

F₂, F₃——株高与粒重为主要组分,反映了生态型内的可变性。

④因子得分与聚类:

表12

水肥组各品种因子得分表

	F ₁	F ₂	F ₃	F ₄	F ₅	F ₆
1	31.7269	54.7565	82.0630	-12.5732	11.1550	8.0220
2	30.3104	49.0454	86.6635	-9.4612	12.5690	8.2625
3	4.1363	45.8511	72.0474	-7.6037	13.7445	8.1849
4	29.5544	53.0924	73.8549	-9.6981	11.5215	8.07844
5	36.4240	44.0573	78.0259	-9.6444	11.1430	7.7780
6	14.9598	39.3506	77.6891	-9.5261	12.6868	7.8087
7	31.5663	38.2534	72.9331	-5.7555	10.9412	8.8733
8	17.8830	45.4754	78.9176	-8.4351	12.1881	7.5492
9	27.4439	72.7962	94.6134	-14.7076	12.2353	7.5643
10	12.1671	62.7314	80.8558	-11.8634	13.6030	8.4854
11	22.5973	50.7977	70.8055	-9.6092	11.6816	8.2311

表13 D²表 (F₁、F₂、F₃得分参加运算)

	1	2	4	11	8	6	5	7	10	9	3
1	0	53.4	74.0	223.5	282.2	528.5	146.5	345.9	452.5	512.2	935.6
2		0	229.4	314.0	227.2	410.2	136.4	305.1	550.2	635.5	908.9
4			0	63.0	219.9	416.5	146.2	225.1	444.2	823.6	701.8
11				0	116.4	236.8	288.8	242.2	352.2	1074.5	366.8
8					0	47.6	346.6	275.2	334.2	1084.2	236.2
6						0	483.0	299.6	564.5	1560.9	191.2
5							0	83.2	945.1	1181.7	1081.5
7								0	1038.3	1680.2	810.9
10									0	523.9	427.0
9										0	1778.7
3											0

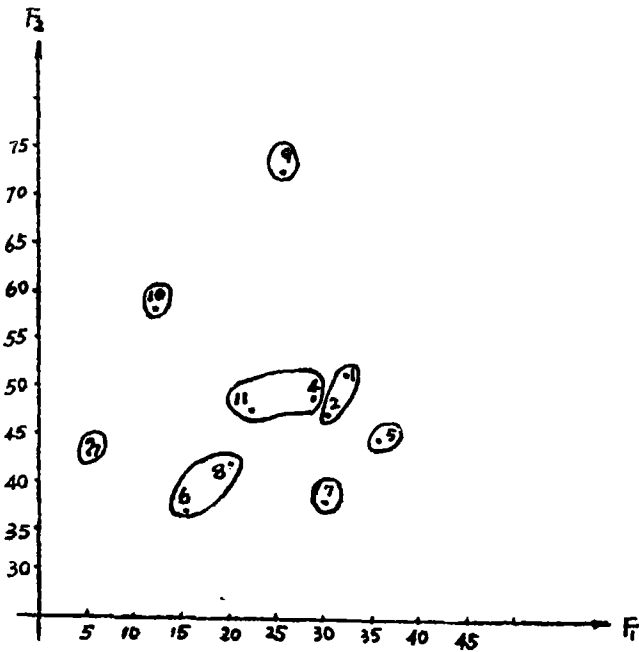


图3 因子得分图

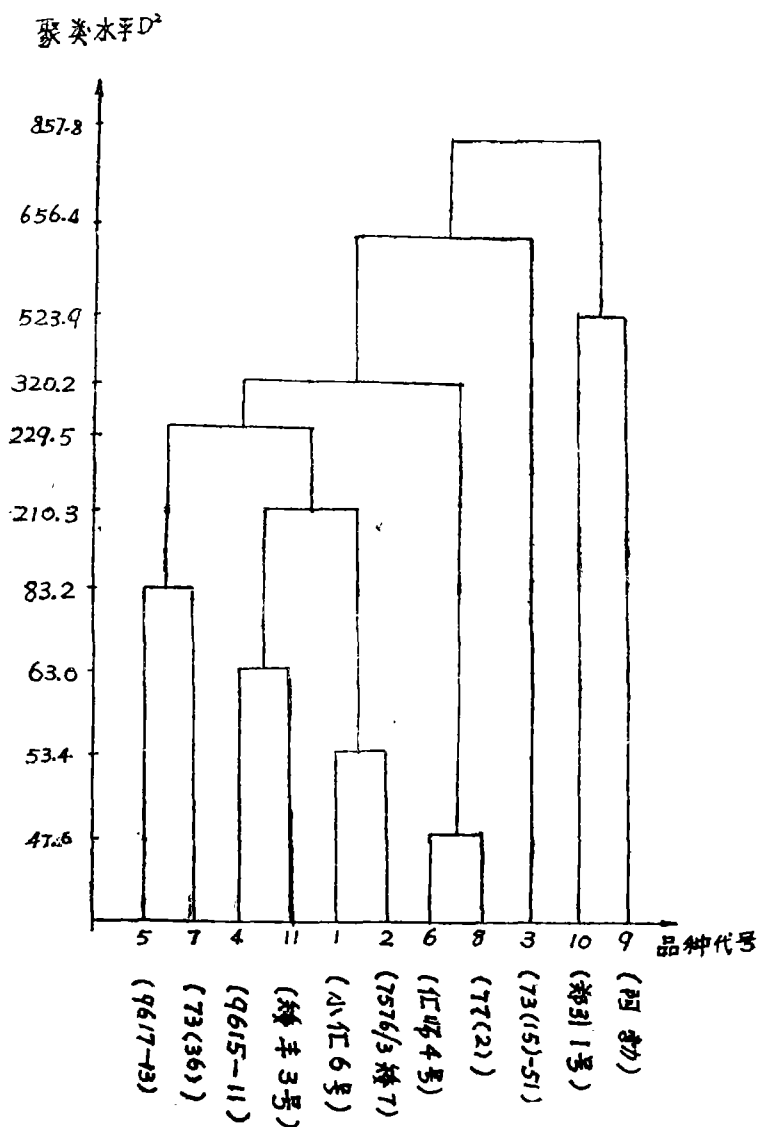


图4 品种聚类图

在聚类水平 $D^2 = 63.0$ 时分成八类

- | | |
|--------------|------------|
| I: {5} | 类内距离为 0 |
| II: {7} | 类内距离为 0 |
| III: {4, 11} | 类内距离为 63.0 |
| IV: {1, 2} | 类内距离为 53.4 |
| V: {9} | 类内距离为 0 |
| VI: {10} | 类内距离为 0 |
| VII: {6, 8} | 类内距离为 47.6 |
| VIII: {3} | 类内距离为 0 |

表14

类间距离表

	I	II	III	IV	V	VI	VII
II	83.2						
III	217.5	233.7					
IV	141.5	325.2	210.3				
V	1181.7	1680.2	949.1	573.9			
VI	945.1	1038.2	398.2	501.4	523.9		
VII	414.8	287.4	246.2	362.0	1322.6	449.4	
VIII	1081.5	810.9	543.3	922.3	1778.3	427.0	213.7

表15

水肥组品种各类性状的组合规律

类别与品种	性状	冬季分蘖	每穗粒数	抽穗期 (4月)	成熟期 (6月)	株高	千粒重
I	5	多	少	20	7	中矮	大
II	7	中	中少	21	7	矮	大
III	4	多	中少	21	8	中矮	小
	11	中	中多	19	7	中矮	小
IV	1	多	中少	21	5	中	中
	2	中	中	18	7	中矮	重
V	9	中少	中	20	7	高	中
VI	10	中少	多	16	5	中	小
VII	6	中少	中	11	2	中矮	大
	8	少	中	13	5	中	大
VIII	3	少	多	10	4	矮	中
按类序性状动态		↓ 多 ↓ 少	↑ 少 ↑ 多	↓ 晚 ↓ 早	↓ 晚 ↓ 早		

2.Q型因子分析

参加试验的11个品种相关系数在0.9604~0.9985之间, 相关矩阵的第一个特征根

$\lambda_1 = 10.8573$, 占方差贡献的99%, 说明它们属于一个大的生态类型。 λ_1 相应的特征向量按大小次序排列以及它们所对应的品种为:

分 量	0.30304	0.30293	0.30285	0.30271	0.30221
品种代号	4	1	8	10	11
0.30137	0.30136	0.30059	0.30056	0.30054	0.29894
6	7	2	5	3	9

由因子得分图及性状的组合规律知, 5 (9615-13), 3 (73(15)5-1) 在分蘖、粒数、早熟性上各居一端, 9 (阿勃) 在株高上很高, 它们所对应的特征分量在小的—端, 说明为极端品种。品种10 (郑引一号), 11 (矮丰3号) 居中, 实际上为代表品种; 4、1、8 在一些性状上有所突破: 1 (小偃6号) 在分蘖上与粒重的组合上趋于代表品种的上限, 8 (77(2)) 在粒重与早熟上趋于代表品种的上限, 但这两个品种都未脱离本地区生态型的规范。

从上述分析中可以看出, 因子分析比单因素分析法所获得的信息全面而集中。

根据上述单因素或多因素分析方法, 可对参试品种的下列组合 $K_0, K_0 + K_1, K_0 + K_1 + K_2, \dots, K_0 + K_1 + \dots + K_i, K_0 + K_1 + \dots + K_i + K_{i+1}$, 逐次进行分析, 便可分析出那些性状组合在生态条件不断变化中是保持稳定的, 那些性状在随着栽培条件的变化而变, 从而看出生态条件在作物品种的改良中所起的制约作用, 这将有助于我们在生态学和遗传学的背景下去解决育种中的决策问题。

五、结果与讨论

陕西关中地区小麦品种生态类型的研究, 前人多采用性状观察比较, 结合生态条件特点进行综合分析的方法。在生态研究中, 这是一种基本的研究分析方法, 具有许多突出的优点; 但它也有一定的局限性, 主要是不易确切的定性定量的分析品种生态型的特征性状。我们研究的主要目的是, 试图探求一种适宜的统计分析方法, 并以此抛砖引玉, 引起争论, 以助此项研究的深入和普遍展开。

(一) 本试验结果表明 (表9、表12), 供试各品种 (系) 在六个主因子上的得分相当一致, 且各分量十分接近, 可以认为属于一个大的品种生态类型。供试品种 (系) 是同一个大生态地区——黄淮麦区大面积种植的品种和选育的稳定优系, 属于同一品种生态类型, 是符合育种和生产实际的。这说明本研究采用的统计分析是基本可行的。

(二) 供试各品种的六个主因子得分图和聚类分析的结果表明, 它们虽属一个品种生态类型, 但仍可划分为若干个生态亚型: 旱肥组共9个品种 (系), 划为四个生态亚型; 水肥组共11个品种 (系), 划为八个生态亚型。由于试验研究的种种局限, 这里划分的生态亚型, 当然不能作为定论。它的意义在于: ①说明品种生态类型是一个具有多

级多层性的开放系统。关中地区的小麦品种生态类型,包含着若干生态亚型。这与前人的研究结果有相似的趋势。这一特点,在研究育种目标的针对性和品种布局等方面应予以充分重视;②说明本研究采用的统计分析方法具有相当的“灵敏”性,在品种生态类型和生态亚型的研究中可以试用。

(三)由因子得分图(图1、图3)可以看出,旱肥组9个品种属于一个类型,其特征为 F_1 轴得正分、 F_2 轴得负分;水肥组11个品种亦属于一个类型,其特征为 F_1 轴得正分、 F_2 轴亦得正分,这说明水、旱两组品种的性状表现的差异相当明显。由于水、旱两组品种的大多数是重复的,仅是水、肥管理条件不同,所以本研究中两组品种在生态分类上的差异,不能认为完全是品种类型的差异。它部分地反映了品种类型的分化(如阿勃、郑引一号与丰产3号、小偃6号),部分地反映了水、肥栽培条件对品种性状表现的强烈影响。这再次表明,育种试验地的地力培养、因地用种和因种栽培的重要性。

(四)用两种方法研究分析的6个性状中,冬季分蘖、每穗粒数和成熟期三个性状较为稳定,可作为关中地区小麦品种生态类型的基本特征;抽穗期、株高和千粒重三个性状变幅较大,可作为可变特征。从性状发育的特点和关中小麦品种性状演变的有关资料看,这一结果有作为进一步研究的参考价值。同时应当指出,用上述6个性状研究品种生态类型的基本特征和可变特征,显然是不够的,这里只是作为品种生态类型研究方法的一种探索而已。

品种生态类型的问题相当复杂,目前还缺乏完善的研究方法,本文仅是一个初步的尝试,有关问题尚需继续研究。

主要参考文献

- 〔1〕赵洪璋、王玉成:《陕西小麦》,1947年,陕西出版社,第41页。
- 〔2〕金善宝等:“中国小麦的种类及其分布”,《南京农学院科学研究专刊》第2号,1959年10月,第34页。
- 〔3〕西北农学院主编:《作物育种学》,1981年1月,第250页。
- 〔4〕赵洪璋等:“小麦杂交育种工作中的若干问题”,《陕西农业科学》,1981年第3期,第3页。
- 〔5〕李振声:“小麦育种工作中几个问题的探讨”,《陕西农业科学》,1981年第3期,第10页。
- 〔6〕中国科学院数学研究所统计组:《方差分析》,1977年,科学出版社,第115页。
- 〔7〕张尧庭、方开泰:《多元统计分析引论》,1982年,科学出版社,第322页。

Research on Statistical and Analytical Method of Ecotypes of wheat varieties and Their Evolution

Yuan Zhifa Song Zhemin Liu Guangzu

(The Northwestern College of Agriculture)

Abstract

The basic characters and evolution laws of ecotypes of wheat varieties are of great significance in a series of basic breeding key links such as working out the breeding objectives/parent material matching/selection of hybrid descendants/identification of accommodation of wheat varieties. As a result, research on ecotypes of wheat varieties has increasingly arrested great attention from the breeders.

Based on research by the predecessors, this paper defined the basic characters of ecotypes and variable meanings and presented the primary statistical method of classification---univariant analysis and multivariate analysis.

Univariant analysis is to use the variance analysis and multiple comparison to judge the differential capacity of performances of wheat varieties so as to determine the stability of performances of them. The quantitative differential types of their performances were given by means of Nair level group method.

Multivariate analysis is to use R-type factor to judge the basic characters and variable natures of ecotypes of wheat varieties. The first component of correlation matrix stands for the basic characters of ecotypes of wheat varieties. The second and third components stand for the variable natures of ecotypes. The subecotypes were differentiated in the terms of accumulation of major factors. The representative species in the eco-environment were analysed and determined by Q-type factors. Information obtained by multivariate analysis is richer, more concentrated and easier to do classification and explanation than that by univariant analysis.