Copyright is owned by the Author of the thesis. Permission is given for a copy to be downloaded by an individual for the purpose of research and private study only. The thesis may not be reproduced elsewhere without the permission of the Author.

EVALUATION OF VARIABILITY IN A FOG GRASS

(HOLCUS spp.) GENE POOL

by

SIAN HOCK TEOW

A thesis

presented in partial fulfilment of the requirement for the

Master

of

Agricultural Science

Massey University

Palmerston North

New Zealand

1978

ABSTRACT

The utimate aim of a breeder's working collection is to utilize the genetic variation for breeding new cultivars. Before this variation can be utilised, it is necessary to obtain the description of attribute from the collection. These can either be obtained from the records of genetic resources (base collection or active collection) or obtained directly from the working collection itself. This will resulted in a huge amount of data. To be of any value, this information need to be classified systematically, and the classification need high degree of objectivity, especially for species of no <u>a priori</u> knowledge.

A Working collection of 160 Yorkshire Fog seed populations, from all over New Zealand, were planted out in Massey University. This formed the gene pool. During Summer 1975, 11 agronomic and morpho-____ logic characters were scored in a semi-quantitative scale. This yield approximately 42,000 data records. These were then systematically reduced to 550 by a series of multivariate analysis techniques. The procedures of Multivariate Analysis of Variance, Multiple Discriminant Analysis and Cluster Analysis were reviewed and their computer programmes were developed.

The clustering behaviours of seven agglomerative polythetic strategies were studied and compared, using the full set of characters. Most of the results concurred with studies carried out by other workers The Minimum Increment Sum of Squares strategy was found to be most suitable for this analysis. A probabilistic decision method was devised to decide objectively, the truncating point for clustering.

For all set of data, the studies did not reveal any ecotypes and hence did not agree wirh the ecoclinal trends hypotheses (of Yorkshire Fog in New Zealand) of Jacquee. The approaches of both studies (of Yorkshire Fog in New Zealand) of Jacques. The approaches of both studies (that of Jacques and the present one) were reviewed critically and a more appropriate approach was suggested for future ecological study.

Preliminary results revealed that there were a few promising groups showing agronomic desirable characters. They were promising breeding materials for future lines selection. Of all the characters studied, flowering date and clump erectness were found to be the most discriminating characters amongst groups, and the most dominant characters in clustering. These implied that selection should be beneficial, if they had moderate high predictive heretability.

ACKNOWLEDGEMENTS

I am deeply indebted to my supervisor, Dr. I. L. Gordon for his guidance and assistance during the course of the work and the preparation of this thesis. Without him this thesis would not have been completed.

Special gratitude must go to my parents, Mr. and Mrs. Teow Lee-Seng, my brothers and sisters for providing me the opportunity to study in New Zealand.

Thanks to Professor B. R. Watkin and all the staff members of Agronomy Department for their assistance and encouragement.

I wish to thank my sisters, Bee Ling, Bee Geok and Bee Ing for their considerable time, effort, patience in typing this thesis. Their help is gratefully acknowledged.

Special thanks are also due to : Dr. D. A. Ratkowsky of CSIRO Hobart, Australia and Dr. R. Pringle of Mathematics Department, Massey University, for their advices on statistical field; Mr. Terry Lynch and the field technicians of Agronomy Department, for maintaining the working collection of Yorkshire Fog; staff members of computer center and library for their assistance; and all those who help in one way or others.

The awards of D. J. McGowan and Helen E. Akers scholarships to partially finance this study are gratefully acknowledged.

TABLE OF CONTENTS

INTRODUCTION

1

CHAPTER 1 LITERATURE REVIEW

1.1	York	shire Fog	3
1.2	Gene	Pool Concept And Maintence	5
1.3	Mult	ivariate Analysis	7
1.4	Assu	mptions for Multivariate Analysis	8
1	.4.1	Multivariate Normal Distribution of Variables	9
1	.4.2	Independence of Observations	10
1	.4.3	Additivity of Effects	10
1	.4.4	Equality of Variance-Covariance Matrices	10
1	.4.5	Transformations	11
1.5	Mult	ivariate Analysis Of Variance (MANOVA)	12
1.6	Muti	ple Discriminant Analysis	14
1.7	Clus	tering Analysis	17
1	.7.1	Similarity Measures	18
1	.7.2	Clustering Procedures	22
1	.7.3	Hierachical Clustering Strategies	24
	1.7	.3.1 Hierachical Divisive	26
	1.7	.3.2 Agglomerative Polythetic Strategies	27
		1.7.3.2.1 Strategies Based on Successive	
		Information Gain	27
		1.7.3.2.2 Strategies Not Based on Successive	
		Information Gain	27
		1.7.3.2.3 Number of Clusters	32

CHAPTER 2 MATERIALS AND METHODS

2.1	Field Design	34
2.2	Field Measurements	34
2.3	Multiveriate Analysis Of Variance (MANOVA) and Multiple	
	Discriminant Analysis (DISCRIM)	37
2	.3.1 MANOVA	37
2	.3.2 Multiple Discriminant Analysis	39
2.4	Similarity Measures	43
2.5	Cluster Analysis	44
2.6	Post Clustering Analysis	44

		Page
CHAPTER 3 R	RESULTS AND ASSOCIATED DISCUSSION	3-
3.1 MAND	DIS	46
3.1.1	ALLCHARA	46
3.1.2	AGROCHARA	55
3.1.3	DISCCHARA	57
3.1.4	JACQCHARA	63
3.1.5	Additional Comments	63
3.2 Comp	parison of Different Clustering Strategies using	
ALLC	CHARA Attributes	67
3.2.1	Results of SEFWIG as Applied to Ward's Method	68
3.2.2	Single Linkage	70
3.2.3	Centroid Method	70
3.2.4	Median Method	70
3.2.5	Average Linkage Between Merged Clusters	73
3.2.6	Average Lingkage Within New-Cluster	73
3.2.7	Complete Lingkage	77
3.2.8	Ward's Method	77
3.2.9	General Comparison	77
3.3 Clus	tering Analysis and Post Clustering Analysis	80
3.3.1	ALLCHARA	80
3.3.2	AGROCHARA	86
3.3.3	DISCCHARA	93
3.3.4	JACQCHARA	96
CHAPTER 4	GENERAL DISCUSSION	
4.1 Mult	ivariate Analysis	104
4.1.1	Multivariate Versus Univariate Analyses	104
4.1.2	Model used	104
4.1.3	Data Transformations	104
4.1.4	Data "Crunching"	105
41.5	Squared Eculidean Distance as a Similarity Measur	е
		105
4.1.6	Probablilistic Decision on Clustering Cut-off	105
4.2 Ecot	ype Studies and Ecoclinal Trends	106
4.3 Agro	nomic and Plant Breeding Aspects	108
4.3.1	Agronomic Relevance of Characters Assessed	108
4.3.2	Limitations of The Study and Subsequent Analyses	109
4.3.3	Variation Amongst Individuals Within Groups	110
4.3.4	Agronomic Evaluation of The Groups (Accessions)	111
4.3.5	Cluster Analysis and The Choice of Parents	111

			Page
CONCLUSION			113
Appendix A-	-1		115
Appendix A-	-2		119
Appendix A-	-3		120
Appendix A	_4		121
Appendix B-	-1		122
Appendix B	-2		132
Appendix C	-1		217
Appendix C	-2		220
BIBLIOGRAPH	Y		221

LIST OF. FIGURES AND TABLES

FIGURES		Page
1.1	The difference between "Group Average Between Merged	
	Clusters" method and "Group Average Within New Cluster"	
	method.	29
3.1	The group centroids ordinated in the 1st (Y_{I}) and 2nd	
	(Y _{II}) discriminant functions.	54
3.2	Changes in SS, and F-test for ALLCHARA, as examined by	
	program SEFWIG.	69
3.3	Dendrogram of ALLCHARA by Single Linkage Method.	71
3.4	Dendrogram of ALLCHARA by Centroid Method, reversals	
	are shown in dotted lines.	72
3.5	Dendrogram of ALLCHARA by Median Method, reversals	
	are shown in dotted lines.	74
3.6	Dendrogram of ALLCHARA by Average Linkage Between Merged	
	Clusters Method.	75
3.7	Dendrogram of ALLCHARA by Average Linkage Within New	76
	Cluster Method.	
3.8	Dendrogram of ALLCHARA by Complete Linkage Method.	78
3.9	Dendrogram of ALLCHARA by Ward's Nethod.	79
3.10	Changes in SS, and F-test for AROCHARA, as examined by	
	program SEFWIG	87
3.11	Dendrogram of AGROCHARA by Ward's Nethod.	88
3.12	Dendrogram of DISCCHARA by Ward's Method	94
3.13	Dendrogram of JACQCHARA by Ward's Method	100

TABLES

3.1	Analyses summary of all characters for 160 groups.	
	(For F-ratio df ₁ = 159, df ₂ = 3643).	47
3.2	The Results of Equality Test of MSCP Matrices, for ALLCHARA	.48
3.2	Results From Multiple Discriminant Analysis, For ALLCHARA.	50
3.4	The S Matrix - The Correlations Between The Discriminant	
	Functions and The Original Characters, for ALLCHARA.	52
3.5	The B Matrix - The Coefficient For Producing Standardized	
	Discriminant Function (Y) from group deviation vectors	
- 8	$(X - \overline{X})$, for ALLCHARA	53

TABLES

3.6 The Results of Equality Test of MSCP Matrices, for AGROCHARA. 56 3.7 The Results From Multiple Discriminant Analysis For AGROCHARA 56 3.8 The S Matrix - The Correlations Between The Discriminant Functions and The Original Characters, for AGROCHARA. 58 3.9 The B Matrix - The Coefficients For Producing Standardized Discriminant Functions (Y) From Group Deviation Vectors $(X - \overline{X})$ for AGROCHARA. 59 3.10 The Results Of Equality Test of MSCP Matrices for DISCCHARA 59 3.11 The Results From Multiple Discriminant Analysis For DISCCHARA. 61 3.12 The S Matrix - The Correlations Between The Discriminant Functions and The Original Characters For DISCCHARA. 61 3.13 The B Matrix - The coefficients For Producing Standardized Discriminant Function from Group Deviation Vectors for DISCCHARA. 62 3.14 The Results Of Equality Test Of MSCP Natrices for JACQCHARA.64 3.15 The Results From Multiple Discriminant Analysis for JACQCHARA. 64 3.16 The S Matrix - The Correlations Between The Functions and The Original Characters For JACQCHARA. 65 3.17 The B Matrix - The Coefficients For Producing Standardized Discriminant Functions From Group Deviation Vectors. 65 3.18 The Proportion of Sum of Sauares (WSS/TSS) Not Explained By Clustering, At Different Stages of Clusstering By Ward's Method for ALLCHARA. 82 3.19 The Ranks of The Means of Each Character of 50 Clusters In ALLCHARA Analysis. 84 3.20 Brief Grouping of The 50 Clusters of ALLCHARA and Their Approximate Average Ranking In Dominant Characters. 85 3.21 The Proportion of Sum of Squares (WSS/TSS) Not Explained By Clustering At Different Stages of Clustering By Ward's Method for AGROCHARA. 89 3.22 The Rank of The Means of Each Character of 44 Clusters In AGROCHARA Analysis. 90 3,23 Brief Grouping of The 44 Clusters of AROCHARA and Their

Approximate Average RAnking In Dominant Characters.

Page

92

TABLES

3.24	The Proportion of Sum of Squares (WSS/TSS) Not Explained	
	By Clustering At Different Stages of Clustering by Ward's	
	Method for DISCCHARA.	95
3.25	The Ranks of The Means of Each Character of 46 Clusters	
	In DISCCHARA Analysis.	97
3.26	Brief Grouping of The 46 Clusters of DISCCHARA and Their	
	Approximate Average Ranking In Dominant Characters.	98
3.27	The Proportion of Sum of Squares (WSS/TSS) Not Explained	
	By Clustering At Different STages of Clustering By Ward's	
	Method For JACQCHARA.	101
3.28	The Ranks of Means of Each Character of 44 Clusters In	
	JACQCHARA Analysis.	102
3.29	Brief Grouping of The 44 Groups of JACQCHARA and Their	
	Approximate Average Ranking In Discriminant Characters.	103

INTRODUCTION

Hill country pasture is important in New Zealand, as it is a highland country, with 28 million acres (64 percent) of the farming land being steep hill country (Scott 1956).

Hill country pasture production can only be maintained if suitable species are grown. In particular, attention must be given to the species' soil fertility requirements. It is common for high producing pasture species to lose their producing ability where soil fertility falls below their requirements, or where practices to suit their best growth are not followed. In the hill country, soil fertility is often poor (White 1973), and species suit to this are preferred.

Yorkshire Fog grass (<u>Holcus lanatus</u> L.), is noted for a virtual absence of edaphic specialization, wide climatic tolerance, low soil fertility requirement, and good adaptation to extensive agricultural (pastoral) system. This grass has established well in New Zealand in humid hill country, unploughable steep hills, acidic peat soils and even swampy lands (Basnyat 1957). It is the major constituent of some 8 million acres of North Island marginal pasture land, of which 5 million acres are in the wetter hill country of the west. It's contribution toward farm productivity has been judged as significant (Munro 1961, Basnyat 1957).

In 1953, Jacques started to investigate Yorkshire Fog as a useful pasture species. The investigation commenced with collecting a wide range of local seed populations from most major areas of New Zealand. After a series of progeny tests and selections, a synthetic cultivar "Massey Basyn" was bred in 1960 (Jacques 1962). A synthetic cultivar is made up of genotypes which have previously been tested for their ability to produce superior progeny when crossed in all combinations (i.e. they have good combining abilities) (Allard 1960).

In a sward productivity trial, Munro (1961) found that "Massey Basyn" compared favourably with perennial ryegrass and performed much better than a commercial Fog line. Riveros(1963) found

1

that dry matter yields were always significantly higher in Yorkshire Fog swards, than in ryegrass swards. In a trial from 1961 to 1964, Watkin and Robinson(1974) Found that "Massey Basyn" had a similar dry matter yield to the ryegrasses (Ariki, Manawa and Ruanui); and the seasonal production of "Massey Basny" well distributed, with relatively good production in winter and summer.

To enable further detailed investigation, and to improve further the agronomic value of Yorkshire Fog grass, Dr. R. G. Clements organized another collection of Fog accessions in 1972. A total of 201 seed populations were collected. These included 108 from the North Island, 89 from the South Island, 3 from Australia and 1 from Spain. This collection was sampled from most parts of New Zealand, even the Westland region (Which Jacques had missed out), and the Northland region (which Munro(1961) and Jacques(1962, 1974) considered as having the most potential for highly productive genes).

Out of the 201 seed populations collected, 160 were planted out as a working collection at Massey University. According to Vavilov, "collection and classification" is the first phase of scientific plant breeding (Frankel 1947); and the aim of the presant study was to examine the phenotypic variability in, and to reveal the relationship amongst, major characters within this collection. In this study, populations have been grouped into phenotypically-similar clusters. The clusters so obtained could be used as sources for breeding material. Several multivariate analysis techniques have been used to achieve this. Firstly, multivariate analysis of variance has been used to investigate whether there were any differences amongst populations. Secondly, multiple discriminate analysis has been used to estimate the relationships between the phenotypic characters, and to ordinate the original scores into uncorrelated discriminant scores. Finally, cluster analysis has been used to group together populations with similar phenotypic patterns of dispersion.