REF. NO.2 MARCH 1989

No.

EXPLORATION AND COLLECTION OF PLANT GENETIC RESOURCES

PART I SEED-PROPAGATED CROPS

TECHNICAL ASSISTANCE ACTIVITIES FOR GENETIC RESOURCES PROJECTS

JAPAN INTERNATIONAL COOPERATION AGENCY

AF
JR
89-42

REF. NO.2 MARCH 1989

EXPLORATION AND COLLECTION OF PLANT GENETIC RESOURCES

PART I SEED-PROPAGATED CROPS

TECHNICAL ASSISTANCE ACTIVITIES FOR GENETIC RESOURCES PROJECTS

JAPAN INTERNATIONAL COOPERATION AGENCY

CONTENTS

Introduction

The Japan International Cooperation Agency (JICA) has sponsored a group training course on plant genetic resources since 1983, and provided some countries with facilities for preserving plant genetic resources as a part of grant aid programs of the Government of Japan. On the basis of such experience, JICA initiated a project for providing further technical assistance to existing genebanks as well as those being planned. As a part of the Technical Assistance Activity for Genetic Resources Projects, GRP newsletter has be published and distributed since last year. In addition to it, it was envisaged to publish a series of manuals, which may be of help for workers in the area of plant genetic resources. Obviously, there are already a number of good references, many of which have been distributed by IBPGR. As it may not make sense to issue similar text books, the new series of manuals is to stress reporting practical experiences of Japanese workers rather than dealing with principles or general concepts. The first of this series has been issued as a collection of papers on the conservation of plant genetic resources.

As the second of this series, this manual deals with exploration or collection aspects of plant genetic resources. In the general part, three papers are collected. Firstly, an introductory paper is provided by Prof. M. TANAKA, a pioneer of plant exploration in Japan, whose experience ranged Asian as well as Latin American regions. Secondly, theoretical aspects for conserving genetic diversity are discussed by Prof. K.YONEZAWA. Thirdly, practical aspects of plant exploration are detailed by Drs. T. NAGAMINE and M. NAKAGAHRA who have led the Laboratory of Plant Exploration at the National Institute of Agrobiological Resources and organized a series of both domestic and international exploration. In the part of individual papers, three reports on domestic collection trips are collected: Dr. M. YAMADA's on domestic maize varieties, a paper by Dr. OKA *et al.* on miscellaneous crops in remote mountainous zones and one by Dr. H. YOSHIKAWA *et al.* on native cruciferous crops in Kinki region (near an old capital). All of these papers are interesting documents to indicate how serious the genetic erosion is in Japan. Last paper by Prof. A. UJIHARA reports preliminary evaluation of buckwheat which was collected by an IBPGR mission in Nepal.

Hiroshi IKEHASHI
Prof. of Plant Breeding, Fac. Horticulture,
Chiba University.

I. Theories and Practices

I –1. Principles and Practices in Exploration and Collection of Plant Genetic Resources

by

Masatake TANAKA

Kihara Memorial Yokohama Foundation

for the Advancement of Life Science

CONTENTS

1) Basic concepts

Civilization is advanced by exploitation of land, and inevitably accompanied by extinction of native organism. The development of agriculture has isolated many organisms for human use from those in the realm of nature, and disrupted the pathes of organism towards accumulation of variation, which depends on natural hybridization and other interactions among organisms. Along with this, so-called "land races" of cultivated plants indigenous to a region have been replaced with modern and uniform ones.

(1) Genetic uniformity and genetic erosion

From the viewpoint of genetic resources, the extinction of native varieties implies the loss of a huge amount of genes. The current tendency of relying on a limited number of cultivars is mainly the consequence of market requirements for high yield, suitability for mechanical harvesting or processing. To some extent, this tendency gathers through technological progress in handling genetic variation, particularly in recent years.

The genetic uniformity in certain crops proceeds in parallel with the evolution of pathogenes virulent to it, which sometimes threaten crop cultivation. Despite of numerous historical lessons that proved the danger of dependence on a limited number of cultivars, genetic bases of certain crops are yet narrowed to a limited number of cultivars. And the trend is aggravated even in core areas of the genetic diversity of cultivated plants.

For instance, in wheat producing areas in the world, the "green revolution" has often driven out land races. The genetic diversity that has been developed in the Middle and Near East, the original place of wheat, has been rapidly reduced to threaten the survival of wheat groups.

The loss of genetic diversity through replacement of diverse native varieties by a few improved varieties is called "genetic erosion". It should not be denied that researchers' effort for varietal improvement and extension of better varieties is one of the major factors for the genetic erosion.

(2) VAVILOV's contribution

The Russian botanist VAVILOV was the first to have established scientific bases in the field of plant exploration. As a pioneer in the research of genetic

resource, VAVILOV surveyed diverse forms for the most of important cultivated plants, and studied their distribution on the earth, while leading the work towards the comprehensive collection of plant materials from all over the world. In this way, he succeeded in discovering native habitats for most of cultivated plants, and reached to a conclusion that a limited areas of rich diversity or gene center can be indicated as the center of the origin of each crop.

Thus, VAVILOV proposed the eight centers of origin of crops, which are now familiar, as shown in Fig.1, and assigned lists of plants to each center as given in Table 1(VAVILOV 1951).

One of VAVILOV's generalizations concerning variation of cultivated plant was what he described as "geographical regularities in the forms of cultivated plants"(VAVILOV 1922). By this, he meant that a similar variation in two or more unrelated crops could be observed in a given area. To systematize many examples of the parallelism, VAVILOV established what he called the Law of Homologous Series, which he claimed also has an predictive value. This law states that similar variation can be found in unrelated crops in a same geographical area.

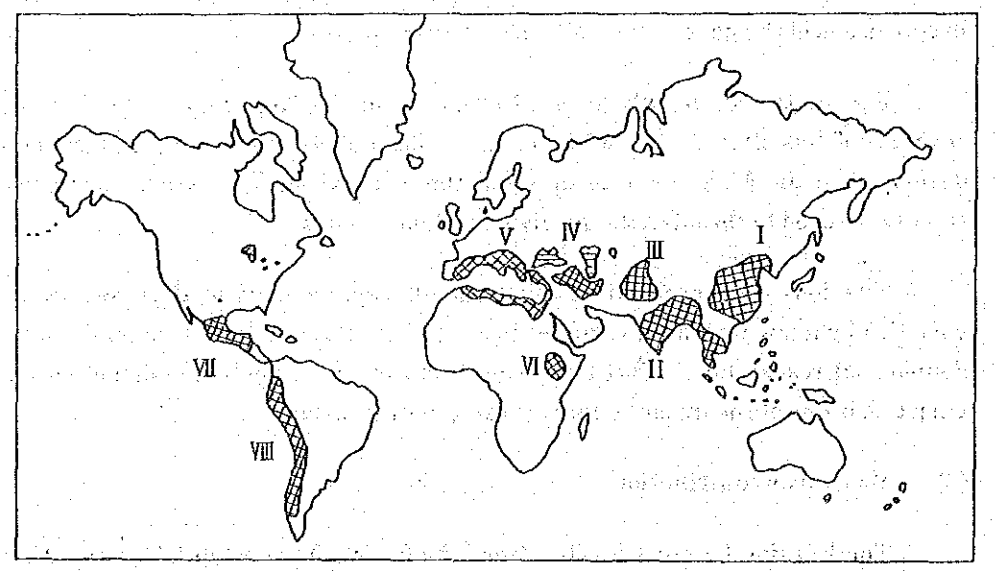


***Fig.1 VAVILOV's eight centers of origin of cultivated plants* (VAVILOV 1951)**

Table 1 World center of origin of cultivated plants (Arranged from VAVILOV 1951)

I	The Chinese Center	
	Azuki bean	*Phaseolus angularis*
	Buckwheat	*Fagopyrum esculentum*
	Chinese tea	*Camellia sinensis*
	Peach	*Prunus persica*
	Pet-sei	*Brassica pekinensis*
	Soybean	*Glycine max*
II	The Indian Center	
	Cucumber	*Cucumis sativus*
	Egg plant	*Solanum melongena*
	Rice	*Oryza sativa*
	Taro yam	*Colocasia antiquorum*
II a	The Indo-Malayan Center	
	Banana	*Musa paradisiaca*
	Bread fruit	*Artocarpus communis*
	Coco-nut	*Cocos nucifera*
	Sugar cane	*Saccharum officinarum*
III	The Central Asiatic Center	
	Apple	*Malus pumila*
	Broad bean	*Vicia faba*
	Grapevine	*Vitis vinifera*
	Onion	*Allium cepa*
	Pear	*Pyrus communis*
	Pistachio nut	*Pistacea vera*
	Radish	*Raphanus sativus*
	Spinach	*Spinacia oleracea*
IV	The Near Eastern Center	
	Barley	*Hordeum vulgare*
	Bread wheat	*Triticum aestivum*
	Carrot	*Daucus carota*
	Macaroni wheat	*Triticum durum*
	Oat	*Avena sativa*
	Rye	*Secale cereale*
V	The Mediterranean Center	
	Asparagus	*Asparagus officinalis*
	Beat	*Beta vulgaris*
	Cabbage	*Brassica oleracea*
	Chick pea	*Cicer arietinum*
	Flax	*Linum usitatissimum*
	Lettuce	*Lactuca sativa*
	Olive	*Olea europaea*
	Pea	*Pisum sativum*
VI	The Abyssinian Center	
	Coffee	*Coffea arabica*
	Okra	*Hibiscus esculentus*
	Sesame	*Sesamum indicum*
	Sorgo	*Sorghum bicolor*
	Teff	*Eragrostis abyssinica*
VII	The South Mexican and Central American Center	
	Common bean	*Phaseolus vulgaris*
	Corn	*Zea mays*
	Pepper	*Capsicum annuum*
	Pumpkin	*Cucurbita pepo*
	Squash	*Cucurbita moschata*
	Sweet potato	*Ipomoea batatas*
	Upland cotton	*Gossypium hirsutum*
VIII	The South American Center	
	Lima bean	*Phaseolus lunatus*
	Pepper	*Capsicum chinense, C. baccatum*
	Potato	*Solanum tuberosum*
	Sea island cotton	*Gossypium barbadense*
	Tabacco	*Nicotiana tabacum*
	Tomato	*Lycopersicum esculentum*
	Winter squash	*Cucurbita maximus*
VIII a	The Chiloe Center	
	Pine strawberry	*Fragoria chiloensis*
VIII b	The Brazilian-Paraguayan Center	
	Cassava	*Manihot utilissima*
	Peanut	*Arachis hipogaea*
	Pineapple	*Ananas comosus*

2) Definition, objects and purposes of exploration

In the primitive ages, human beings obtained their food merely by picking or hunting, mostly by chance without any insight into its nature. Later, in the stone ages, they collected food with the clear aim to select useful plants for their life. This is a kind of exploration, and as a consequence useful plants were improved through cultivation towards the creation of cultivated plants.

(1) Definition of exploration

Strictly speaking, "exploration" means to look for something with a specific aim and an intention to utilize it. To look for something unknown without any definite object in mind beforehand is called "expedition". On the other hand, "collection" premises preservation not necessarily implying specific purposes or utilization. Therefore, collection is often used in contrast to exploration.

Where man copes with genetic resource problems, the ideal should be to deal with them under the concept of "exploration". VAVILOV's book (1951) gave us a general theory relating to exploration. In future, we must develop better methodology to find desired objects from nature in deliberate consideration.

(2) Objects in exploration

Generally, the objects of exploration are grouped into wild species and cultivated species. Exploration of wild species may include following categories:

a. Development of new genetic resources such as plants for hydrocarbons,

b. Utilization of wild relative species as a genetic pool such as disease resistance genes,

c. Reconstitution of plants by the use of ancestral species.
Exploration of cultivated species, especially land races may includes the following categories:

d. Re-evaluation of plants for new characteristics, for example, that of wheat for lysine and protein contents,

e. Accumulation of genetic resources through their preservation to expand genetic diversity.

(3) Purposes of exploration

FRANKEL and SOULE (1981) classified genetic resources to be explored into two groups according to purposes as follows:

a. Materials which are now on the verge of extinction or feared to be so in near future. This exploration is regarded as a kind of relief work, but it is of great importance in the light of the present world situation. The objects include cultivars which are rapidly driven away by the introduction of improved varieties, as has seen dramatically shown in the "green revolution" and in many cases of crop replacement for economic reasons. Recently, the genetic erosion caused by improved cultivars is aggravated, resulting in the rapid reduction of genetic diversity. Although the significance of the preservation of land races is well understood, in reality it tends to be overlooked for several reasons.

b. Materials which are worthy for researches and plant improvements currently underway or planned for future.

This kind of exploration is on the line to the original objective of the activity for genetic resources. The main subject in this case includes both wild species, particularly ancestral wild species concerned with creation of cultivated plants, and allied species. Further, all plants which are included in gene pools for cultivated plants may be considerred. Thus, this category includes the exploration for the utilization through re-evaluation of characteristics of wild species and existing cultivars. As mentioned, this exploration aims at their utilization through which new resources might be developed.

(4) Classification of gene pools

HARLAN and DE WET (1971) proposed the gene pool classification in order to treat several cereal crops on a uniform basis (Table 2). In this system the total array of variation within maximum genetic reach is partitioned into primary, secondary, and tertiary gene pools.

The primary gene pool includes all that can be crossed with a given crop species, yielding reasonably fertile hybrids in which chromosomes pair well and

Table 2 Primary and secondary gene pools of the major cereals. (HARLAN and DE WET 1971)

Crop	Ploidy Level	PRIMARY GENE POOL			SECONDARY GENE POOL
			Spontaneous Species		
		Cultivated Species	Wild Races	Weed Races	
Einkorn	2×	*Triticum monococcum*	*T. boeoticum*	*T. boeoticum*	*Triticum, Secale, Aegilops*
Emmer	4×	*T. dicoccum*	*T. dicoccoides*	None	" " "
Timopheevi	4×	*T. timopheevi*	*T. araraticum*	None	" " "
Bread Wheat	6×	*T. aestivum*	None	None	" " "
Rye	2×	*Secale cereale*	*S. cereale*	*S. cereale*	" " "
Barley	2×	*Hordeum vulgare*	*H. spontaneum*	*H. spontaneum*	None
Oats	2×	*Avena strigosa*	*A. hirtula; A. wiestii*	*A. strigosa*	*Avena* spp.
Ethiopian Oats*	4×	*A. abyssinica; A. vaviloviana*	*A. barbata*	*A. barbata*	"
Oats	6×	*A. sativa*	*A. sterilis*	*A. sterilis; A. fatua*	"
Rice	2×	*Oryza sativa*	*O. rufipogon*	*O. rufipogon*	*Oryza* spp.
African Rice	2×	*O. glaberrima*	*O. barthii*	*O. stapfii*	"
Sorghum	2×	*Sorghum bicolor*	*S. bicolor*	*S. bicolor*	*S. halepense*
Pearl Millet	2×	*Pennisetum americanum*	*P. violaceum*	*P. americanum*	*P. purpureum*
Maize	2×	*Zea mays*	*Z. mexicana*	*Z. mexicana*	*Z. perennis Tripsacum* spp.

* Hardly more than encouraged weeds.

in whose offspring genetic segregation is reasonably normal. The primary gene pool corresponds to the widely accepted concept of the biological species. The primary gene pool of a crop almost always includes wild and weedy races as well as cultivated ones.

The secondary gene pool includes all the species that can be crossed with a given crop species but with restricted gene flow. Genes can be transferred from the secondary to the primary gene pool, but one must struggle with those barriers that separate biological species, such as sterility, poor chromosome pairing, lethal or weak hybrids, or poorly adapted hybrid derivatives and so on.

The tertiary gene pool includes all the species that can be crossed with a given crop, but the hybrids are not easily obtained. The hybrids may be lethal or completely sterile. If any gene transfer is possible at all, it must be through a radical manipulation of the reproduction system such as embryo culture, tissue culture, use of complex hybrid bridges and so on.

3) Methods of exploration and collection

Practical methods comprise collections at research organizations such as experimental stations and universities as well as at markets, depending on purposes. However, the most important is the exploration or collection in natural habitats and cultivated fields. To choose target regions, such factors as natural habit of specific plants to be explored, the center of diversity of cultivated plants, and the ecological environments for specific genes have to be taken into consideration. For effective exploration and collection, the range of variation of individual traits of the plants as well as taxonomic background has to be studied beforehand. Furthermore, it is desirable to collect as many specimens as possible, even if they belong to one species as far as they grow in different regions.

(1) Exploration in natural habitat areas of specific plants

As VAVILOV (1951) stated in the theory of gene center, in topographically isolated mountainous areas such as Caucasus and Ethiopia, an accumulation of variation can be seen, but even in such regions exploration of natural habitants is substantially difficult. In general, we have little chance to find such a field as a natural habitat, unless we are informed by local people. In other words, such fields as native habitat are found mostly by chance. Nevertheless, if we step into

the site after studying aimed plants from an ecological viewpoint, some specific sites out of entire projected regions may come out as promising. In arid zones such as deserts, for instance, burrow pits, those sites where in rainy season flooding forms a temporary river should be the place to be explored. Also, a river can offer a significant route for distribution. In particular, such area where a river bends and flooded water often covers a broad basin, a native habitat of wild species can be found. It is noteworthy that seed is distributed by rivers bringing about a wide range of distribution of the species.

Generally, species related to the origin of cultivated plants introgressed into cultivated fields as companion plants, and have been reproduced for generations as contaminant seed. For example, *Aegilops squarrosa*, the ancestor of bread wheat, is often present as a companion plant in wheat fields. As agricultural development proceeds, a land is converted to cultivated field, pastures are grazed, and the population becomes denser, the native fields of wild species are inevitably narrowed. In such areas as sanctuaries protected for special purposes, we often find that a species on the verge of extinction is still surviving in a native habitat. For instance, historical remains which people have long preserved at every cost can be good places for exploration.

(2) Exploration in fields of cultivated plants

Collection in the fields which are not yet dominated by modern cultivars is the best possible way to gather extensively useful genes as well as variants of a given crop. In this case, it is recommendable to collect as much variants as possible in one field. It enables us to cover a variation of physiological traits which are often invisible. Although a random sampling might be conceivable, a more effective way is to collect as much variation as possible in terms of visible morphological characters in the occasion where the purpose is to collect genetic resources.

(3) Market collection

Market collection can offer chances of collecting information and materials of land races in projected regions. A wide range of variation and heterogeneity is not necessarily found in under-developed state of a region. In some regions, it is frequently seen that pure varieties have been introduced more recently. In other words, the genetic diversity only tells us the history of the cultivation, especially in such a region, where we can collect buried useful genes.

In a market collection, it is necessary to confirm where commodities of a given market is mainly shipped from. Sometimes, we can collect useful genetic resources in a small rural town more efficiently than in a big city.

4) Record taking in exploration and collection

Precise description of collection *in situ* is important above all. Insufficient records about sites decreases its value tremendously. We would expect that records are precise enough to enable ones to visit the spot for re-collection.

My recording method is cited from my field note as follows:

(Example) June 10. First stop at a site where the meter of car shows 1,584km from A City (driving scale of 1,520km at the central plaza, a landmark). Second stop at a site of 1,592km.

Field note description should be as folows:

1,520 the Central plaza of A City
1,584 6-10-1 right side of the road
elevation 1,200m
a. *Aegilops squarrosa* black spike
b. *Aegilops squarrosa* yellow spike
c. *Ae. ovata*

1,592 6-10-2 15m outside the left road
elevation 1,100m
a. *Triticum boeoticum*
b. Triticeae, species is unknown
c. *T. durum* (wheat field)

The advantage of this system are; a collection site is clear and the sign on the bag, 5-10-1a, helps subsequent arrangement.

In exploration in a foreign country appropriate procedures for the transportation of collected materials to a mother country have to be finished before departure in accordance with plant quarantine law. As some of countries ban carrying out of plants, thorough examination and consideration are

required.

5) Arrangement of collected materials

Collected materials are to be examined systematically from a variety of aspects. In doing this, particularly careful consideration should be given to the first seeding of collected materials. In fact, this is a very difficult task. In order to conserve and maintain a wide range of genetic resources, bulk harvesting is desirable, while in order to conserve specific characters individual selection is also necessary. Here, I would like to propose the followings for a guideline; a) a stock seed has to be preserved perpetually, b) the seed of the first generation should be harvested separately for each plant. Seed from all the plants, or otherwise as many as possible, should also be preserved. The reason for these two ways comes from a fear that collected materials tend to be handled based on researcher's personal evaluation, and collected materials other than those receiving special attention tend to be ignored and lost.

Collected materials by a plant exploration should be conserved perpetually at any efforts. For instance, establishing a national seed conservation center will contribute very much to the conservation of valuable materials. Also, a government should make great efforts to help the conservation and preservation of them.

6) *Some examples of exploration for the plant genetic resources*

(1) Rust tolerance in the *Aegilops squarrosa* L.

Collection using ecological characteristics as indices, such as habitats under various environmental conditions, is expected to lead to the exploration of physiologically useful genes. Consequently, we can often find, as expected, such valuable genes as for disease or cold resistance.

Our bread wheat, namely 6x wheats was originated as an amphidiploid of a hybrid between a species of emmer wheat and *Aegilops squarrosa*. It has been known that *Ae. squarrosa* is distributed in the vast area of east of Caucasus, namely Eastern Turkey, Iran, Afghanistan, Pakistan, Armenia, Azerbaijan and Georgia.

After EIG(1929) this species is classified into two subspecies, namely ssp. *eu-squarrosa* EIG and ssp. *strangulata* EIG (Fig.2).

Ssp. *eu-squarrosa* contains three varieties, *typica*, *meyeri* and *anathera*.

A series of field survey and collection of wheats and their closely related genus *Aegilops* were conducted in Pakistan, Afghanistan and Iran by a team of researchers of the Kyoto University Scientific Expedition to the Karakoram and Hindukush, in 1955 (KIHARA *et al.* 1965)(Fig.3).

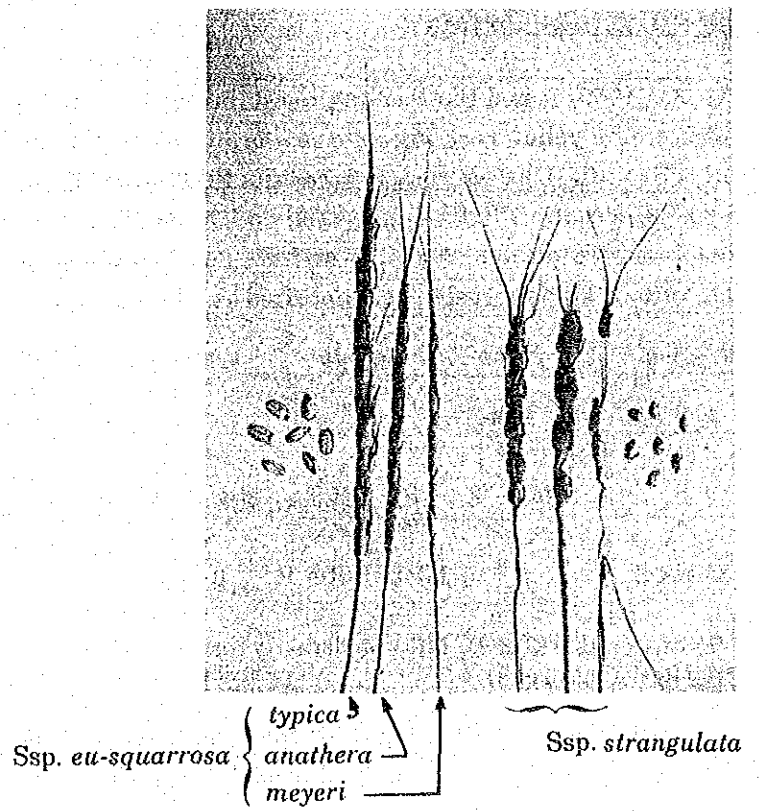


Fig.2 Ears of *Ae. squarrosa*

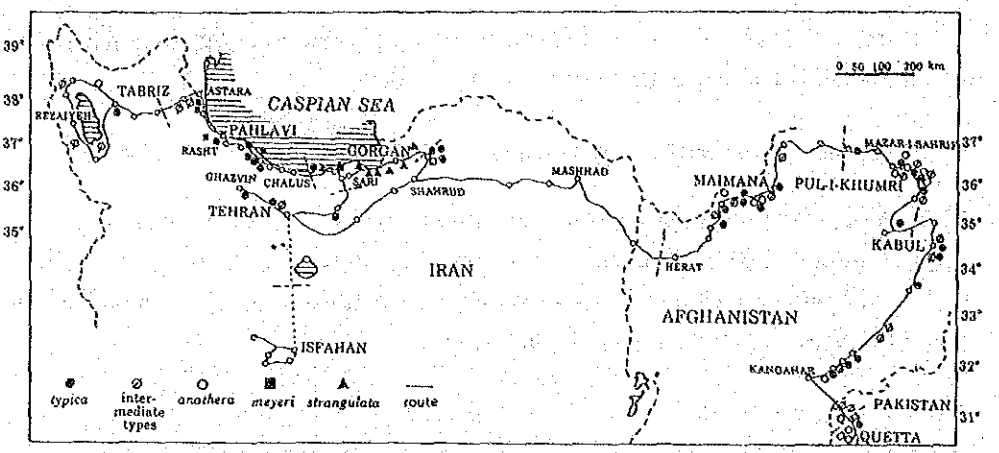


Fig. 3 Map showing the localities of *Ae. squarrosa* along the route of Kyoto Univ. Sci. Expedition 1955

Dr. N. HIRATSUKA (1959) tested the seedling resistance of 136 strains of *Ae.squarrosa* to uredospores of yellow rust, *Puccinia striiformis* WESTENDORF (*P. glumarum* ERIKSSON *et* HENNING), brown rust, *P. recondita* ROBERGE *et* DESM. f. sp. *tritici* and black rust, *P. graminis* PERS, f. sp. *tritici*, and divided the types of susceptibility to rust into five groups, A,B,C,D and E. The relationship of susceptibility types to various kinds of rust are summarized as follows:

	A-group	B-group	C-group	D-group	E-group
yellow rust	S	S MR-S	S	S	MR-S
brown rust	S	R-MR R	S MR-S	MR R-MR R	S
black rust	S MR-S	S	MR or R-Rm R	R-MR R	S

Remarks: S=Susceptible (3 or 4). MR=Moderately resistant (2).
R=Resistant (0 or 1).

According to his experimental data, almost all the strains of var. *typica*, var. *anathera* and the intermediate type, are susceptible to all three kinds of rust, while all strains of var. *meyeri* and most of ssp. *strangulata* are resistant to yellow rust.

Namely, as shown in Table 3, all strains of var. *meyeri* were resistant to the brown and black rusts, being of C- or D-type resistance, 11 out of 16 strains of ssp. *strangulata* were also resistant to the same kinds of rust.

Table 3 Susceptibility of *Ae. squarrosa* to rusts (KIHARA et. al. 1965)

Variety	Susceptibility					Total
	A-group	B-group	C-group	D-group	E-group	
typica	73(83.9%)	5(5.7%)	7(8.0%)	1(1.2%)	1(1.2%)	87
anathera	10(90.9%)	1(9.1%)	0	0	0	11
intermediate type between *typica* and *anathera*	13(86.9%)	0	2(13.3%)	0	0	15
strangulata	3(18.8%)	2(12.5%)	9(56.2%)	2(12.5%)	0	16
meyeri	0	0	2(50.0%)	2(50.0%)	0	4
intermediate type between *typica* and *strangulata*	0	0	2(66.6%)	1(33.4%)	0	3
Total	99	8	22	6	1	136

Table 4 Frequency of *squarrosa* strains and varieties in eight regions, with respect to susceptibility to rusts (HIRATSUKA 1959)

Region	Varienty	Susceptibility					Total
		A-group	B-group	C-group	D-group	E-group	
Quetta	*typica*	3	0	0	0	0	7
	anathera	4	0	0	0	0	
Kabul	*typica*	20	0	0	0	0	25
	anathera	1	0	0	0	0	
	intermadiate type[1]	4	0	0	0	0	
Pul-i-Khumri	*typica*	16	0	0	0	1	24
	anathera	4	0	0	0	0	
	intermediate type[1]	3	0	0	0	0	
Maimana	*typica*	17	0	0	0	0	21
	anathera	1	0	0	0	0	
	intermediate type[1]	3	0	0	0	0	
Tehran	*typica*	7	2	1	0	0	12
	anathera	0	0	0	0	0	
	intermediate type[1]	1	1	0	0	0	
Gorgan	*strangulata*	3	2	9	2	0	19
	intermediate type[2]	0	0	2	1	0	
Pahlavi	*typica*	4	2	6	1	0	19
	meyeri	0	0	2	2	0	
	intermediate type[1]	0	0	2	0	0	
Tabriz	*typica*	6	1	0	0	0	9
	intermediate type[1]	2	0	0	0	0	
Total		99	8	22	6	1	136

1) Intermediate type between var. *typica* and var. *anathera*.
2) Intermediate type between var. *typica* and ssp. *strangulata*.

However, a few strains of var. *typica* and the intermediate type showed indications of C- or D-type resistance. Table 4 summarizes the types of susceptibility to various kinds of rust of a number of strains and varieties of *Ae. squarrosa* from eight regions. The majority of the strains with C- or D-type resistance were collected in the Gorgan and Pahlavi regions. It is interesting that var. *typica* from the Pahlavi regions is resistant, though the var. *typica* from other regions is susceptible. Namely, among the 13 strains of var. *typica* collected in the Pahlavi region, seven strains had C- or D-type resistance.

Since the climate is warm and humid in the Gorgan and Pahlavi regions along the Caspian coast, resistant varieties or strains would have been brought into existence in those regions by natural selection.

The seedlings of F_1 hybrids between the resistant and susceptible strains and their progeny were inoculated with uredospores of *Puccinia graminis* f. sp. *tritici* and *P. recondita* f. sp. *tritici*. In the cross, *meyeri*×*typica*, the resistance of *meyeri* was dominant. The F_2 segregation seemed to be monogenic (Table 5).

On the basis of genomic and morphological analyses, KIHARA(1949) established the genealogical relationship in wheat and *Aegilops*. It was found that the genus *Triticum* consists of three basic genomes, A,B and D, and our dinkel or bread wheat (hexaploid wheat, AABBDD) was originated as an amphidiploid of a hybrid between a species of emmer wheat (tetraploid wheat, AABB) and *Ae. squarrosa* (DD). Since it was made sure that *Ae. squarrosa* is one of ancestors of dinkel wheat, the synthesis of allohexaploid wheats from various crosses of emmer wheats with *Ae. squarrosa* has been attempted on a wide scale.

Table 5 Segregation of resistance vs. susceptibility in the varietal crosses in *Ae. squarrosa* (HIRATSUKA 1957)

Parent and cross combination		*P. recondita* f. sp. *tritici* 21B	*P. graminis* f. sp. *tritici* 21
var. *meyeri* (No. 2144-1)		R	R-MR
var. *typica* (No. 1)*		MR-S	S
var. *typica* (No. 2107-4)		S	S
var. *meyeri* (No. 2144-1) × var. *typica* (No. 2107-4)	F_1	R	R
	F_2	R(23) S(3)	R(18) S(7)
var. *meyeri* (No. 2144-1) × var. *typica* (No. 1)*	F_1	R	R
	F_2	R(16) S(5)	R(18) S(4)

S: Susceptible, MR: Moderately resistant, R: resistant.
*: Strain from the old collection of Kyoto University.

Table 6 Susceptibility of *Ae. squarrosa*, *Triticum* species and their amphidiploids to rusts (HIRATSUKA, unpub).

Parental species or amphidiploid	Susceptibility to: Yellow rust *Puccinia striformis* (*P. glumarum*)	Brown rust *Puccinia recondita* f. sp. *tritici* 21B	Black rust *Puccinia graminis* f. sp. *tritici* 21
Ae. squarrosa strangulata No. 2112	S	R-MR	S
" " " No. 2118	S	R-MR	S
" " *meyeri* No. 2114	S	R	R-MR
T. dicoccum (Vernal)	S	S	S
T. durum (Gulab)	S	S	S
T. persicum stramineum	S	S	S
ABD 13, *T. dicoccum* (Vernal) ×*Ae. squarrosa strangulata* (No. 2112)	S	S	S
ABD 14, *T. durum* (Gulab) ×*Ae. squarrosa strangulata* (No. 2118)	S	S	S
ABD 16, *T. durum* (Gulab) ×*Ae. squarrosa meyeri* (No. 2114)	S	MR-S	MR-S
ABD 18, *T. persicum stramineum* ×*Ae. squarrosa strangulata* (No. 2112)	S	S	S
ABD 19, *T. persicum stramineum* ×*Ae. squarrosa strangulata* (No. 2118)	S	S	S
ABD 22, *T. persicum stramineum* ×*Ae. squarrosa meyeri* (No. 2114)	S	R	S

From Table 6, we see that a synthesized 6x wheat strain, ABD 22, was resistant to brown rust; it was obtained from the cross, susceptible *persicum* var. *stramineum*×resistant *meyeri*. However ABD 16, which possesses the same resistant *meyeri* genomes, was not resistant to brown rust, and it was classified as moderately resistant to susceptible. Also, in this case the emmer species was susceptible to brown rust. So the inheritance pattern is apparently not simple.

As to the relationship to black rust, ABD 22, an amphidiploid between susceptible *persicum* var. *stramineum* and resistant *meyeri*, was susceptible. In this case the susceptible character of *persicum* is dominant (or epistatic) over the resistant character of *meyeri*. So far as our investigations are concerned, all amphidiploids derived from susceptible parents were susceptible. So it seems that resistant 6x strains can be synthesized only when the *squarrosa* parent has the resistant gene.

Accordingly, using these strains, a new disease-resistant bread wheat has been synthesized successfully.

(2) Salinity tolerance in wheat and *Aegilops*

One-third of the world's land is in arid soil areas. These arid soil areas have been partly utilized through the aid of irrigation, and are partially productive. In the arid areas, even if enough water is availabe by usual irrigation systems, this will induce salt accumulation and result in sterile soils, as many past examples have shown.

Salinity tolerance is becoming of ever greater importance in the wheat cultivation of the Mediterranean region, which are arid zones. The following *Triticum* and *Aegilops* strains were evaluated for their degree of salt-tolerance: Tetraploid species (10 strains of *T. durum*; 2 strains of *Ae. ovata;* 7 strais of *Ae. variabilis*; 3 strains of *Ae.kotschyi*) and hexaploid species (2 strains of *T.aestivum*).

A 1/2,000 diluted Hyponex solution was used to fertilizers all sand cultures. The plants were examined for their tolerance to 0 ppm, 5,000 ppm and 10,000 ppm solutions of NaCl. Estimation was made of plant height, number of productive tillers per plant, number of spikelets and grains per ear, number of grains per plant and 100-kernel-weight. Some important data are shown in Table 7 (TANAKA and EGAWA, 1981). Two strains of *T.aestivum* (KU 264) and

Table 7 Salt tolerance-test in *Triticum* and *Aegilops* (TANAKA and EGAWA, 1981)

Species	(Strain No.)	Na Cl-ppm	Length	Number of fertile ear	Number of rachis/ear	Number of grains/ear	Number of grains/ear	Weight/ 100 grains
Triticum durum	(KU3668)	5,000	73	63	102	89	–	20
		10,000	71	75	95	56	–	19
	(KU3674)	5,000	72	63	86	89	–	24
		10,000	72	75	86	74	–	26
T. aestivum	(KU264)	5,000	76	80	89	80	98	55
		10,000	85	70	93	94	98	49
	(KU3846)	5,000	55	67	65	13	21	–
		10,000	50	67	30	0	0	–
Aegilops ovata	(KU6006)	5,000	75	66	97	90	59	43
		10,000	76	69	97	97	67	33
	(KU6023)	5,000	56	56	102	0	0	–
		10,000	58	58	100	0	0	–
Ae. variabilis	(KU6649)	5,000	70	67	–	57	–	27
		10,000	75	72	–	55	–	28
	(KU6675)	5,000	76	74	–	82	–	26
		10,000	74	83	–	68	–	29
Control		0	100	100	100	100	100	100

Ae. ovata (KU 6006) showed high tolerance viz. even in a 10,000 ppm NaCl solution 100-kernel-weight was reduced to only ca. 1/3-1/2 compared with the control, number of grains per ear and/or plant did not differ from that of the control. It is noteworthy that a saline-tolerant strain was found in the cultivated strains of *T.aestivum* (common wheat). KU 264 (Shirasagi-komugi) is a popular wheat variety cultivated in the Chugoku region of Japan. It should be possible to use this strain as a saline-tolerant variety in saline soil areas, and also to use it as a saline-tolerant gene source in wheat breeding.

Genotypic variation has been reported to affect the amount of homoeologous pairing in hybrids with polyploid wheats in *Ae. speltoides* (DVORÁK 1972; KIMBER and ATHWAL 1972). As shown in Table 8, the mean chromosome pairing at M1 in *Ae. speltoides* (SS, 2n=14) × tetraploid wheat (AABB, 2n=28) hybrid was large. The range of mean univalent frequencies of these hybrids was 2.99 to 8.00 per cell (TANAKA and SANO, 1980). The amount of homoeologous pairing observed seems to depend on the *speltoides* genotypes used. The strain 7712, for example, showed a high level of pairing in hybrids, while 7716 showed a low level. No apparent difference due to the genomic constitution of the wheat parent was observed. *Ae. speltoides* suppresses the activity of the gene *Ph* in wheat chromosome 5BL and thus allows not only homologous but also homoeologous chromosome pairing to occur in hybrids (RILEY *et al.* 1961).

Table 8 Mean chromosome pairing at M1 in *Aegilops speltoides* (n=7) × tetraploid wheat (n=14) hybrid (TANAKA and SANO, 1980)

Ae. speltoides × tetraploid wheat	Univalent	Bivalent	Trivalent
5715 × 114	4.80	4.18	2.56
7712 × 110	2.99	3.46	3.58
7716 × 125	8.00	3.83	1.76
7719 × 125	6.18	4.40	1.98
7727 × 114	5.45	5.27	1.64
7730 × 125	5.35	3.89	2.56
7731 × 125	4.62	3.65	2.97
7757B × 125	5.58	4.30	2.26
7770 × 114	4.90	4.42	2.42
7786 × 125	3.87	3.73	3.14
7800 × 114	5.12	4.28	2.44
7848 × 110	5.60	4.13	2.31

Therefore, a useful way to transfer saline-tolerant genes from an allied species to wheat for promising wheat breeding in saline soil areas may be to use a convenient gene of *Ae. speltoides* which facilitates chromosome pairing between incomplete homologous chromosomes. As shown in Table 8, strain KU 7712 greatly facilitated the chromosome pairing compared with other *speltoides* strains.

Strain KU 6006 of *Ae. ovata*, collected from the Dead Sea area, may be a practical gene source for transfer of the saline-tolerant gene of *Aegilops* spp. to cultivated wheat varieties of *T. aestivum*. *T. aestivum* (2n=42) are crossed with *Ae. ovata* (2n=28) to produce 35-chromosome hybrids. The hybrids which make successive back-cross to *T.aestivum*, and some individuals with 2n=44 (22_{II}) in the pedigree will be selected. When these individuals are crossed with *Ae. speltoides* such as strains KU 7712, promising saline-tolerant plants may be bred in the offspring with 2n=42 (21_{II})(Fig.4).

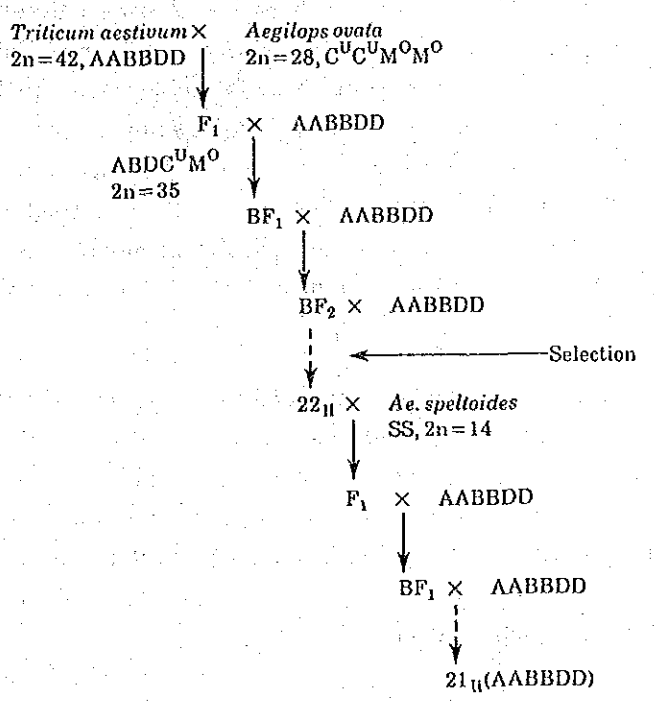


Fig.4 Breeding scheme in the introduction into wheat of the salinity tolerance of *Aegilops ovata*

RILEY *et al.*(1968) has already succeeded in introducing the yellow rust resistance of *Ae. comosa* (MM, 2n=14) into *T. aestivum* by the same means.

(3) Variation of flowering time with teosinte

The close relative of maize(*Zea mays* L.), teosinte, has been assumed to be a plant important in elucidating the evolution of maize under domestication. All investigators are in agreement that maize and teosinte are closely related, and that they have a common source of germ-plasm in a phase of their evolution. In addition, teosinte is an important crop of the pasture at present.

Teosintes were collected from eight regions, namely Iguala, Chilpancingo, Amecameca, Morelia, Cd. Guzman, Manantlan and Nobogame in Mexico, and El Progreso in Guatemala by the members of the Kyoto University Scientific Exploration to Mexico and Guatemala in 1981-1982 (TANAKA 1983). Also, Nobogame teosinte was obtained from the collection of Dr.T. A. K. KATO, Colegio de Postgraduados ENA, Chapingo. The materials from each region were divided into seven groups according to Iltis-Doebley's and Wilkes's systems. Both teosintes collected in the Iguala and Chilpancingo regions of Mexico can be referred to as *Z. mays* ssp. *parviglumis* var. *parviglumis* (Balsas race). Teosintes collected from Amecameca, Morelia and Nobogame regions of Mexico can be referred to as *Z.mays* ssp. *mexicana* and were further identified respectively as Chalco, Central plateau and Nobogame race. Teosinte collected from El Progreso region of Guatemala can be referred to as *Z. luxurians* and were identified as the Guatemala race.

A perennial type of teosinte was collected from two regions, namely Cd. Guzman and Manantlan. The former was identified as *Z.perennis*, namely a tetraploid type, and the latter as *Z. diploperennis*, namely a diploid type. Some progenies of the samples of perennial teosinte collected in two the regions were raised in Kyoto in 1982, and studied cytologically. As a result, their chromosome numbers were ascertained respectively as 4x and 2x. And, 4x and 2x plants were not mixed in each of their original populations.

Table 9 Variation in flowering time of teosinte strains from different localities in Kyoto (TANAKA, 1983)

EMG NO.	Species and variety	Race	Locality or source	Altitude (m)	Latitude (N)	Days to flowering Male	Female
1127	*Zea mays* ssp. *mexicana*	Nobogame	Tarahumare Valley, S. Chihuahua, Mexico	–	ca. 26°	66	75
1119	"	Central plateau	28 km N of Morelia, Michoacan, Mexico	1890	19°50'	107	113
1106	"	Chalco	17 km N of Amecameca, Mexico, Mexico	2240	19°10'	107	112
1101	*Z. mays* ssp. *parviglumis* var. *parviglumis*	Balsas	73 km NW of Iguala, Guerrero, Mexico	1650	18°30'	119	129
1113	*Z. luxurians*	Guatemala	2 km N of El Progreso, Jalapa, Guatemala	1010	14°50'	121	131
1126	*Z. diploperennis*	perennial 2x	1.5km E of La Joyas (Sierra de Manantlan), Jalisco, Mexico	2180	19°35'	118	128
1124	*Z. perennis*	perennial 4x	Piedra Ancha, near Los Depositos (19 km SSW of Guzman), Jalisco, Mexico	2050	19°40'	118	121
Control	*Z. mays* ssp. *mays*	"Aso"	Kumamoto, Japan	–	ca. 35°	75	82

All known varieties of teosinte are short-day plants, though previous investigations have shown that teosinte varieties differ to a considerable degree in their response to photoperiod. Therefore, to observe the response of several varieties, the collected strains were grown under natural daylength in the greenhouse of the Plant Germ-plasm Institute, Kyoto University, and both the flowering time of tassel and ear were observed.

Strains from seven different localities were used in this experiment. Five strains of annual teosinte, 1 strain of diploid perennial teosinte, 1 strain of tetraploid perennial teosinte and 1 strain of a Japanese cultivar of maize were grown in this experiment.

The number of days to flowering of each strain are given in Table 9. In the experiment, the strains from the lower latitude reached the flowering stage later than those from a higher latitude. In addition, it was found that the Nobogame teosinte formed perfect seed, but other teosinte strains produced seeds not fully ripening under the natural daylength in Kyoto.

7) Conclusion

It is a well known fact that many organisms on the earth keep on decreasing with the advance of civilization. Unfortunately, we have to admit the fact that the conservation of stocks of academic value or stocks which are not directly related to economic benefits is neglected with little care for preservation. Wishing real human welfare and well being, we should work urgently for exploring nature where numerous valuable and useful genes are buried, while preserving them.

8) References

DVORÁK, J. 1972. Genetic variability in *Aegilops speltoides* affecting homoeologous pairing in wheat. Can. J. Gent. Cytol. 14, 371-380.

EIG, A. 1929. Monographisch-kritische Ubersicht der Gattung *Aegilops*. Repertorium specierum movarum regni vegetabilis, Beihefte 55.

FRANKEL, O. H. and M. E. SOULE 1981. Conservation and Evolution. Cambridge Univ.Press, pp.327.

HARLAN, J. R. and J. DE WET 1971. Toward a rational classification of cultivated plants. Taxon 20:509-517.

HARLAN, J. R., J. DE WET and E. G. PRICE 1973. Comparative evolution of cereals. Evolution 27:311-325.

HIRATSUKA, N. 1957. Susceptibility of varieties of *Aegilops squarrosa* to black rust (*Puccinia graminis* f.sp. *tritici*) and brown rust (*P. triticina*). Rep.Kihara Inst. Biol. Res. (Seiken Ziho)No.8, 9-10.

HIRATSUKA, N. 1959. Susceptibility of varieties of *Aegilops squarrosa* to yellow rust (*Puccinia striiformis*), brown rust (*P.recondita* f. sp. *tritici*) and black rust (*P. graminis* f. sp. *tritici*). Wheat Information Service Nos.9-10, 34-41.

KIHARA, H. 1949. A new synthesized 6x-Wheat. Proc.8th Intern.Cong. Genet. Hereditas, Suppl. Vol., 307-319.

KIHARA. H., K. YAMASHITA and M. TANAKA 1965. Morphological, physiological, genetical and cytological studies in *Aegilops* and *Triticum* collected from Pakistan, Afghanistan and Iran. In Results of the Kyoto University Scientific Expedition to the Karakoram and Hindukush, 1955. Vol. 1, Kyoto Univ. pp. 1-118.

KIMBER, G. and R. S. ATHWAL 1972. A reassessment of the course of evolution of wheat. Proc. Natl. Acad. Sci. USA 69, 912-915.

RILEY, R., G. KIMBER and V. CHAPMAN 1961. Origin of genetic control of diploid-like behavior of polyploid wheat. J. Hered. 52, 22-25.

RILEY, R., V. CHAPMAN and R. JOHNSON 1968. The incorporation of alien disease resistance in wheat by genetic interference with the regulation of meiotic chromosome synapsis. Genet. Res. 12, 199-219.

TANAKA. M. and J. SANO 1980. Effect of the chromosome pairing gene and B-chromosome in *Aegilops speltoides*. Japan,J.Breeding 30 (Suppl.2),134-135 (in Japanese).

TANAKA, M. and Y. EGAWA 1981. Survey of salt tolerance in *Triticum* and its relatives. Japan J. Breeding 31 (Suppl.2), 22-23 (in Japanese).

TANAKA, M. 1983. Field survey on wild relatives of maize (teosinte) in Mexico and Guatemala. Rep. Plant Germ-plasm Inst. Fac. Agr. Kyoto Univ. 6, 3-8.

VAVILOV, N. I. 1922. The law of homologous series in variation. Jour. Genet. 12, 47-89.

VAVILOV, N. I. 1928. Geographische Genzentren unserer Kulturpflanzen. Verbandlungen des V. Int. Kongr. Vereb. Wissenschaft. Berlin 1927. Aeits. f. Ind. Abst. u. Vererbungl. Suppl. 1, 342-369.

VAVILOV, N. I. 1951. The origin, variation, immunity and breeding of cultivated plant. *Chronica Botanica*. 13, 1-364.

I –2. Appropriate Size of Sample for Collecting Germplasms from Natural Plant Populations

by

Katsuei YONEZAWA

Kyoto Sangyo University

CONTENTS

1) Introduction

Natural vegetations are being rapidly degraded and a world-wide standardization of crop cultivars is progressing in these years (e.g., ECKHOLM 1976, RAMADE 1981), causing a rapid and large-scale extinction of valuable genetic resources. Establishing methodology for conservation of genetic variation is then one of the most urgent problems to be solved for future generations. Genetic variations in natural habitats may be conserved in two ways, i.e., in natural state as a natural genetic reservoir, or, under highly managed conditions. For the latter type of conservation, appropriate methods and techniques for collection and maintenance of germplasms must be developed.

Sampling method for collection of germplasms from natural plant populations is discussed in this report. A field collection project may be carried out on either of two principles, i.e., to collect plants (genotypes) which have some particular (and rare in most cases) phenotypic characteristic, or, to collect an unbiased sample so that the genetic variation of a target population may unbiasedly be taken up in the sample. On the first principle, there is no problem concerned with sampling method because in this case collectors do not do sampling but seek for the plants having the particular characteristic aimed. On the second principle, the result of collection largely depends on the method of sampling.

Three problems must be answered to settle optimal sampling method for field collection. The first is concerned with how to choose or identify the target population (collection sites) to be sampled. A geographical space, plant individuals in which can be regarded as members of one Mendelian population, is decided based on the meteorological, topographical, pedological and biological factors which determine the geographical distribution of plant genotypes and species in the target area (e.g., HAWKES 1983, MORISHIMA *et al.* 1984). The second problem deals with how to choose plants from within a target population. To meet the unbiasedness of sample, plants must be sampled so that the germplasms (alleles and genotypes) collected may be regarded as a random sample from the target population. Completely random sampling, stratified random sampling, and some systematic or non-random sampling (e.g., grid sampling) are thinkable as choices. Comparison of these methods remains to be investigated in relation to the spatial distribution of genotypes within target populations.

The third problem is concerned with the size of sample, i.e., how many plants and seeds are to be sampled from a target population. Investigations made hitherto on sample size are reviewed and developed in this report.

2) Sufficient size of sample from a target population

Appropriate sample size for a target population depends on the genetic constitution of the population and the main concern of collectors. Genetic constitution of target population and the aim of collection are defined in this section, based on which goodness of sampling is formulated and calculated to estimate a sufficiently large size of sample for a target population.

(1) Mathematical model

a. Target population

A population of a diploid plant species is assumed as the target population (site) to be sampled. If there are two alleles, A_1 and A_2, at a locus of interest, three genotypes A_1A_1, A_1A_2 and A_2A_2 are expected to segregate in this target population. Denoting the frequencies of these genotypes by G_1, G_2 and G_3, and assuming that plants in this population are self-pollinating with a selfing rate s, the genotypic array in the seed embryos borne on single plants can be formulated as shown in the upper part of Table 1. When the two alleles are selectively neutral and this population is at genetic equilibrium, the following relations hold between the frequencies of genotypes and alleles (e.g., ALLARD *et al.*1968);

$$\begin{aligned} G_1 &= p_1^2 + p_1 p_2 f \qquad \{f = s/(2-s)\} \\ G_2 &= 2p_1 p_2(1-f) \\ G_3 &= p_2^2 + p_1 p_2 f \end{aligned}$$

where f measures the deviation from random mating.

Table 1. Genotypic arrays in target population and seeds on single plants

Plants in population		Seeds on single plants				
Genotype	Freq.	Genotype	Freq. in selfed seeds	Freq. in outcrossed seeds	Freq. in the total	: Designation
Two alleles:						
A_1A_1	G_1	A_1A_1	1	p_1	$s+(1-s)p_1$	: g_{11}
		A_1A_2	0	p_2	$(1-s)p_2$	: g_{12}
		A_2A_2	0	0	0	
A_1A_2	G_2	A_1A_1	1/4	$p_1/2$	$s/4+(1-s)p_1/2$	: g_{21}
		A_1A_2	1/2	1/2	$s/2+(1-s)/2$	: g_{22}
		A_2A_2	1/4	$p_2/2$	$s/4+(1-s)p_2/2$	: g_{23}
A_2A_2	G_3	A_1A_1	0	0	0	
		A_1A_2	0	p_1	$(1-s)p_1$	: g_{32}
		A_2A_2	1	p_2	$s+(1-s)p_2$	: g_{33}
Three alleles:						
A_1A_1	G_1	A_1A_1	1	p_1	$s+(1-s)p_1$	: g_{11}
		A_1A_2	0	p_2	$(1-s)p_2$	: g_{12}
		A_1A_3	0	p_3	$(1-s)p_3$	: g_{13}
		A_2A_2	0	0	0	
		A_2A_3	0	0	0	
		A_3A_3	0	0	0	
A_1A_2	G_2	A_1A_1	1/4	$p_1/2$	$s/4+(1-s)p_1/2$	: g_{21}
		A_1A_2	1/2	$(p_1+p_2)/2$	$s/2+(1-s)(p_1+p_2)/2$	: g_{22}
		A_1A_3	0	$p_3/2$	$(1-s)p_3/2$	: g_{23}
		A_2A_2	1/4	$p_2/2$	$s/4+(1-s)p_2/2$	: g_{24}
		A_2A_3	0	$p_3/2$	$(1-s)p_3/2$	: g_{25}
		A_3A_3	0	0	0	
A_1A_3	G_3	A_1A_1	1/4	$p_1/2$	$s/4+(1-s)p_1/2$	: g_{31}
		A_1A_2	0	$p_2/2$	$(1-s)p_2/2$	: g_{32}
		A_1A_3	1/2	$(p_1+p_3)/2$	$s/2+(1-s)(p_1+p_3)/2$	: g_{33}
		A_2A_2	0	0	0	
		A_2A_3	0	$p_2/2$	$(1-s)p_2/2$	: g_{35}
		A_3A_3	1/4	$p_3/2$	$s/4+(1-s)p_3/2$	: g_{36}
A_2A_2	G_4	A_1A_1	0	0	0	
		A_1A_2	0	p_1	$(1-s)p_1$	: g_{42}
		A_1A_3	0	0	0	
		A_2A_2	1	p_2	$s+(1-s)p_2$	: g_{44}
		A_2A_3	0	p_3	$(1-s)p_3$	: g_{45}
		A_3A_3	0	0	0	
A_2A_3	G_5	A_1A_1	0	0	0	
		A_1A_2	0	$p_1/2$	$(1-s)p_1/2$	: g_{52}
		A_1A_3	0	$p_1/2$	$(1-s)p_1/2$	: g_{53}
		A_2A_2	1/4	$p_2/2$	$s/4+(1-s)p_2/2$	: g_{54}
		A_2A_3	1/2	$(p_2+p_3)/2$	$s/2+(1-s)(p_2+p_3)/2$	: g_{55}
		A_3A_3	1/4	$p_3/2$	$s/4+(1-s)p_3/2$	: g_{56}
A_3A_3	G_6	A_1A_1	0	0	0	
		A_1A_2	0	0	0	
		A_1A_3	0	p_1	$(1-s)p_1$	: g_{63}
		A_2A_2	0	0	0	
		A_2A_3	0	p_2	$(1-s)p_2$	: g_{65}
		A_3A_3	1	p_3	$s+(1-s)p_3$	: g_{66}

s = selfing rate ; Two alleles : p_1 = frequency of allele A_1 in the target population ($=G_1+G_2/2$) ; p_2 = frequency of allele A_2 ($=G_2/2+G_3$); Three alleles : p_1 = frequency of allele A_1 ($=G_1+(G_2+G_3)/2$) ; p_2 = frequency of allele A_2 ($=+G_4+(G_2+G_5)/2$); p_3 = frequency of allele A_3 ($=G_6+(G_3+G_5)/2$)

Six genotypes will segregate in the population if there are three alleles A_1, A_2 and A_3 at the locus. Definitions and formulations for the three-allele case are presented in the lower part of Table 1. If the population is at genetic equilibrium for three neutral alleles, the following relations are expected to hold,

$$
\begin{aligned}
G_1 &= p_1^2 + p_1(1-p_1)f \\
G_2 &= 2p_1 p_2 (1-f) \\
G_3 &= 2p_1 p_3 (1-f) \\
G_4 &= p_2^2 + p_2(1-p_2)f \\
G_5 &= 2p_2 p_3 (1-f) \\
G_6 &= p_3^2 + p_3(1-p_3)f
\end{aligned}
$$

where f is defined the same as in the two-allele case.

b. Situations

Optimal method of sampling varies with the purpose of sampling. Four situations below are studied in this report;

I_a : All of the alleles contained in a target population are to be collected.

I_b : One or some particular (mostly rare) alleles are to be collected.

II_a : All of the genotypes contained in the target population are to be collected.

II_b : One or some particular genotypes are to be collected.

Sampling in situations I_a and I_b may be referred to as sampling for alleles, since in these cases not genotypes but alleles are the object of collection. Noticing that the genetic potential of a target population can be maintained, or somehow recovered in case of necessity, if all or aimed alleles have been collected, these situations may be assumed in most if not all field exploration projects.

Plants, if not fully selfing, bear many seeds (embryos) with different genotypes. So, to keep the allelic multiplicity of a target population, or to take up particular alleles, not many plants need to be sampled if a sufficiently large number of seed embryos are harvested from each plant. Goodness of sampling in

situations I_a and I_b is formulated in the next section in terms of the numbers of plants and seeds per plant.

In situations II_a and II_b, on the other hand, sampling for genotypes is the issue, i.e., not alleles but genotypes are the object to be collected. Situation II_a may be the case when the collection is made for some research survey, for instance, to infer the frequencies and spatial distribution of genotypes in the target population. Situation II_b may be assumed, for instance, when plants having some valuable characteristic which is not identifiable by outward appearance are to be drawn. In situations II_a and II_b, the genotypic multiplicity of seed embryos on single plants has no concern with goodness of sampling, which therefore is defined only in terms of the number of plants sampled.

c. Criteria to measure goodness of sampling

Four types of probabilities mentioned below, which respectively correspond to the four situations describe above, are used in this report to measure goodness of sampling;

P_{Ia} : the probability of all of the alleles contained in a target population being collected,

P_{Ib} : that of a rarest allele being collected,

P_{IIa} : that of all of the genotypes contained in a target population being collected,

P_{IIb} : that of a rarest genotype being collected.

Assuming that m plants with n seeds for each plant are randomly chosen from a target population as defined in section a, the probability P_{Ia} for the two-allele case is formulated as,

$$P_{Ia} = 1 - (G_1 g_{11}{}^n + G_2 g_{21}{}^n)^m - (G_2 g_{23}{}^n + G_3 g_{33}{}^n)^m$$

where genotype frequencies G_1, g_{11}, etc. follow the definitions in the upper part of Table 1.

The probability for the three-allele case is formulated as

$$P_{Ia} = 1 - P_1 - P_2 - P_3 + P_4 + P_5 + P_6$$

where,

P_1 = the probability that allele A_1 is missed in a total of $N = mn$ seeds sampled,

P_2 = the probability similarly defined as above for allele A_2,

P_3 = the probability similarly defined for allele A_3,

P_4 = the probability that both of alleles A_1 and A_2 are missed, in other words, only alleles A_3 are taken up,

P_5 = the probability similarly defined for alleles A_2 and A_3,

P_6 = the probability similarly defined for alleles A_3 and A_1.

Using the parameters defined in the lower part of Table 1, the above probabilities are calculated by,

$$P_1 = \{G_2(g_{24}+g_{25})^n + G_3(g_{35}+g_{36})^n + G_4(g_{44}+g_{45})^n + G_5(g_{54}+g_{55}+g_{56})^n + G_6(g_{65}+g_{66})^n\}^m ,$$

$$P_2 = \{G_1(g_{11}+g_{13})^n + G_2(g_{21}+g_{23})^n + G_3(g_{31}+g_{33}+g_{36})^n + G_5(g_{53}+g_{56})^n + G_6(g_{63}+g_{66})^n\}^m ,$$

$$P_3 = \{G_1(g_{11}+g_{12})^n + G_2(g_{21}+g_{22}+g_{24})^n + G_3(g_{31}+g_{32})^n + G_4(g_{42}+g_{44})^n + G_5(g_{52}+g_{54})^n\}^m ,$$

$$P_4 = (G_3g_{36}{}^n + G_5g_{56}{}^n + G_6g_{66}{}^n)^m ,$$

$$P_5 = (G_2g_{24}{}^n + G_4g_{44}{}^n + G_5g_{54}{}^n)^m ,$$

$$P_6 = (G_1g_{11}{}^n + G_2g_{21}{}^n + G_3g_{31}{}^n)^m .$$

Assuming alleles A_1 and A_2 in the two-allele case to be common and rare alleles of frequencies p_1 and p_2 respectively, the probability P_{1b} is given by,

$$P_{1b} = 1 - (G_1g_{11}{}^n + G_2g_{21}{}^n)^m$$

It is noted that this formula can be applied also to the three-allele case, if alleles A_1 and A_2 together are assumed to be common and A_3 to be rare (the allelic frequency p_1 now shows the aggregate frequency of A_1 and A_2, and p_2 does

the frequency of A_3). In this report, rare alleles and genotypes stand for the ones which are as frequent as or less frequent than 0.05, unless otherwise stated.

The probability P_{IIa} is formulated by,

$$P_{IIa} = 1 - \sum_{h=1}^{3} (1 - G_h)^m + \sum_{h=1}^{3} G_h^m$$

for the two-allele case, and

$$P_{IIa} = 1 - \sum_{h=1}^{6} (1 - G_h)^m + \sum_{\substack{h,i=1 \\ (h<i)}}^{6} (1 - G_h - G_i)^m - \sum_{\substack{h,i,j=1 \\ (h<i<j)}}^{6} (1 - G_h - G_i - G_j)^m$$
$$+ \sum_{\substack{h,i,j,k=1 \\ (h<i<j<k)}}^{6} (1 - G_h - G_i - G_j - G_k)^m$$
$$- \sum_{\substack{h,i,j,k,l=1 \\ (h<i<j<k<l)}}^{6} (1 - G_h - G_i - G_j - G_k - G_l)^m ,$$

for the three-allele case. The probability P_{IIb} is presented by

$$P_{IIb} = 1 - (1 - G_{min})^m$$

where G_{min} shows the frequency of rarest genotypes.

Calculation of the probabilities considering multiple independent loci is straightforward. The probabilities now are obtained simply by the multiplication of the probabilities which were calculated for each locus using the above mentioned formulae. When, for instance, two loci with allelic frequencies $(p_1, p_2) = (0.95, 0.05)$ and $(p_1, p_2, p_3) = (0.40, 0.30, 0.30)$ are considered, the probability that all of the two and three alleles at the respective loci are collected is obtained by the multiplying the two probabilities calculated for each locus with the respective allelic frequencies. When, for another example, three loci with allelic frequencies $(p_1, p_2, p_3) = (0.90, 0.05, 0.05)$ each are considered, the probability that all of the three alleles at each of the three loci, i.e., nine alleles in total, are collected is given by the cube of the probability calculated for any one locus among the three. This way of calculation applies also to the other three probabilities P_{Ia}, P_{IIa} and P_{IIb}.

(2) Numerical computations

a. Sampling for alleles

The probabilities P_{Ia} and P_{Ib} were calculated for some probable combinations of the allelic frequencies p_1, p_2 and p_3, selfing rate s and sample size parameters m and n. There was no substantial difference between P_{Ia} and P_{Ib}, for any cases calculated, although the former was slightly smaller than the latter, meaning that common alleles have little concern in determining sample size in both of the situations I_a and I_b. Results of computations only for P_{Ia} are mentioned below.

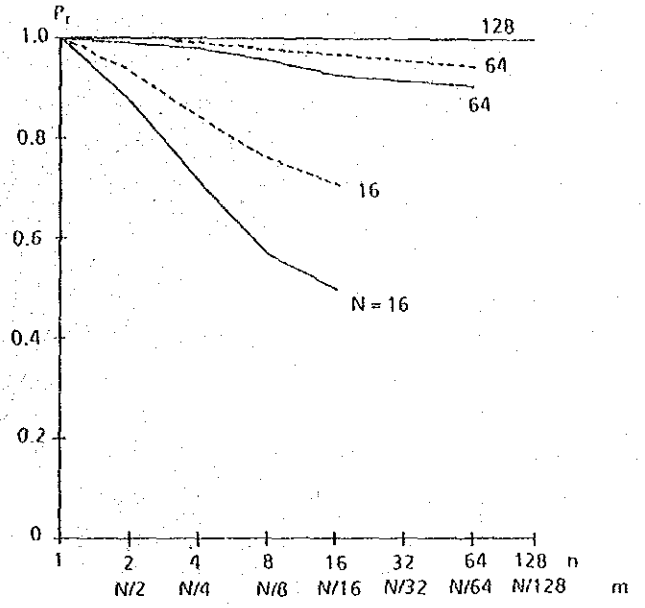
Fig.1 shows the values of P_{Ia} with varying combinations of the numbers m and n under three fixed values (16, 64 and 128) of the total seed number, $N = mn$. In Fig.1, not the absolute values of P_{Ia} but the ratio to the case of $m = N$ and $n = 1$, i.e. P_r, were presented for convenience. It can be seen from this Figure that with both of the two selfing rates assumed ($s = 0.1$ and 0.9), the probability is largest at the combination of $m = N$ and $n = 1$, and decreases steadily with decreasing values of m. This indicates that the number of plants per population rather than that of seeds per plant is the primary determinant of the success of sampling.

The importance of the plant number m, however, differs largely with the rates of selfing. In predominantly outcrossing populations ($s = 0.1$), not a large difference is caused when either m or n is increased, unless N is very small and/or very rare alleles are to be collected. In predominantly selfing populations ($s = 0.9$), on the other hand, the decrease in m causes a sharp reduction in P_r especially when rare alleles (of a frequency of 0.05 or less) are targeted, indicating that the outcome of sampling largely depends on the number of plants sampled.

In most practical projects involving field collection, collectors have to visit many sites with a limited amount of time and labor. The number of plants per site in this situation cannot be freely chosen, but is limited to a relatively small value. Then, increasing the number of seeds per plant is the only choice of collectors since this requires little extra cost.

The effect of increasing the number of seeds per plant with a fixed number of plants can be seen in Fig.2, where the probability P_{Ia} with different values of n is shown. Fig.2 reveals that the merit of increasing the seed number varies with the rate of selfing in the target population. In outcrossing populations, the

a) s = 0.1



b) s = 0.9

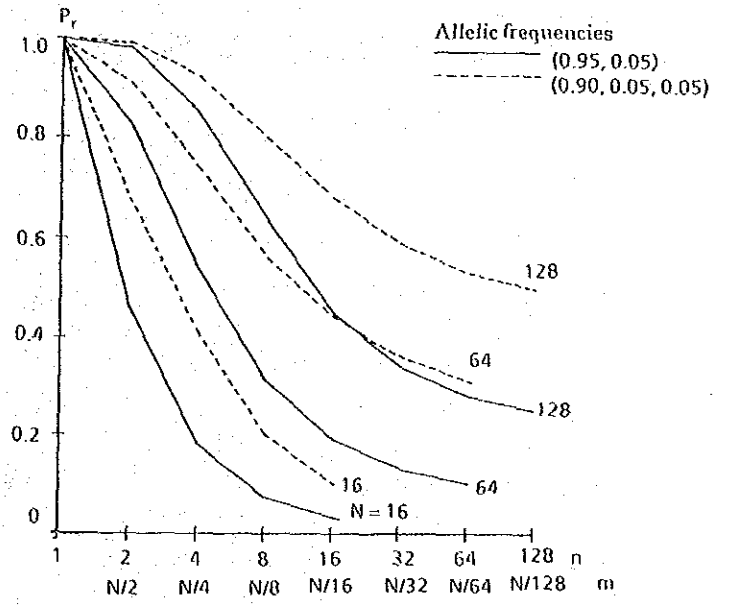
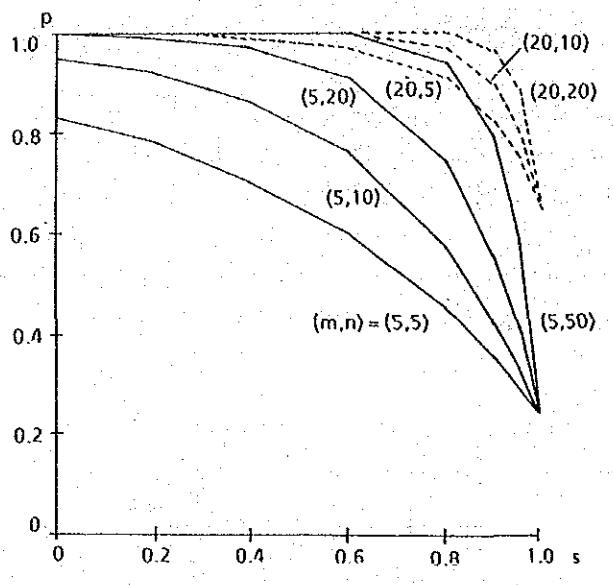


Fig. 1. The probability (relative value to the case of m = N and n = 1) of all segregating alleles being collected in different combinations of the numbers of plants (m) and seeds per plant (n).

a) $(p_1, p_2) = (0.95, 0.05)$



b) $(p_1, p_2, p_3) = (0.95, 0.05, 0.05)$

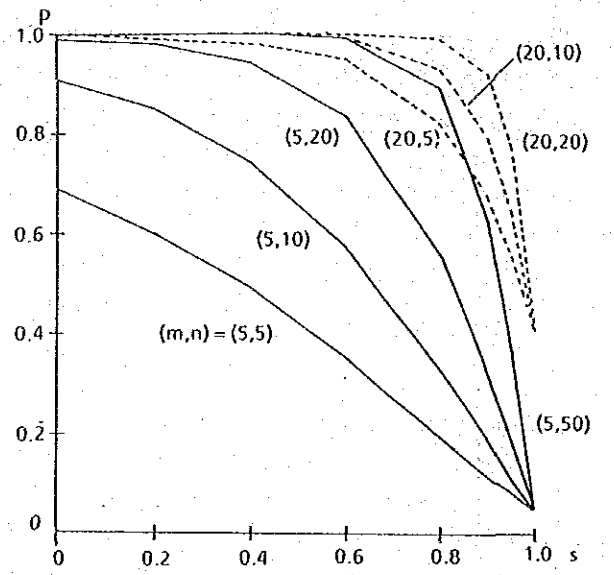


Fig. 2. The probability of all alleles being collected with varying rates of selfing of a target population.

probability P_{Ia} takes a fairly large value even with a relatively small number of seeds, and rapidly gets to a sufficiently high value after a rather small increase of *n*. In completely or highly selfing populations with a selfing rate larger than about 0.95, the probability is not or only a little improved by the increase of n. In any cases calculated, the largest improvement in the probability due to increased seed number per plant is produced in populations with moderately high values of selfing rate ranging from 0.6 to 0.8 approximately.

Table 2 gives the seed number per plant, n, and the total seed number, $N=mn$, which are necessary to give a probability as high as 0.95 for four different values of the plant number m. Computations for the case of multiple independent loci were also presented. The allelic frequencies and selfing rates assumed in Table 2 are considered to be in the range of real situations (e.g., JAIN and MARSHALL 1967, BROWN 1978, OKA and MORISHIMA 1967, VASEK and HARDING 1976). For reference, the genotype arrays in the eighteen different situations assumed for a single locus were shown in Table 3.

It is pointed out from Table 2 that, in any cases calculated, not many plants need to be drawn if a few hundred seeds can be taken from each plant. Only one plant seems to be sufficient in most situations. There is no difference between cases (0.95, 0.05) and (0.50, 0.45, 0.05), and between $(0.95, 0.05)^2$, (0.90, 0.05, 0.05) and $(0.90, 0.05, 0.05)\times(0.40, 0.30, 0.30)$, indicating that the sample size is determined not by the total number of alleles and loci concerned, but by the number of rare alleles whether located at a single locus or different loci.

It should be noted, however, that the sample size does not increase very sharply with increasing number of rare alleles; in Table 2, no large differences occur between cases (0.90, 0.10) and $(0.90, 0.10)^2$, between (0.95, 0.05) and $(0.95, 0.05)^2$, and between (0.90, 0.05, 0.05), $(0.90, 0.05, 0.05)^2$, $(0.90, 0.05, 0.05)^4$. A fuller picture of the contribution of the number of rare alleles can be seen from Fig.3, which shows diagramatically the change in sample size (seed number per plant) with increasing number of loci and rare alleles. It is obvious in this Figure that the increase in sample size due to the increased number of rare alleles is rather small unless the plants are highly selfing.

Table 2. Number of seeds to be collected in order to give a probability (P_{Ia}) as high as 0.95 of taking up all alleles included in a target population.

Allel freq.*	(0.70, 0.30)												(0.70, 0.30)²											
s	0.2				0.5				0.8				0.2				0.5				0.8			
m	1	3	5	10	1	3	5	10	1	3	5	10	1	3	5	10	1	3	5	10	1	3	5	10
n	9	2	1	1	15	3	2	1	42	9	3	1	13	3	2	1	20	5	2	1	53	13	5	1
N	9	6	5	10	15	9	10	10	42	27	15	10	11	9	10	1	20	15	10	1	53	39	25	10

Allel freq.	(0.80, 0.20)												(0.80, 0.20)²											
s	0.2				0.5				0.8				0.2				0.5				0.8			
m	1	3	5	10	1	3	5	10	1	3	5	10	1	3	5	10	1	3	5	10	1	3	5	10
n	15	4	2	1	25	7	3	1	67	18	8	2	19	5	3	1	32	9	4	2	83	23	11	3
N	15	12	10	10	25	21	15	10	67	54	40	20	19	15	15	10	32	27	20	20	83	69	55	30

Allel freq.	(0.90, 0.10)												(0.90, 0.10)²											
s	0.2				0.5				0.8				0.2				0.5				0.8			
m	1	3	5	10	1	3	5	10	1	3	5	10	1	3	5	10	1	3	5	10	1	3	5	10
n	34	10	5	2	56	17	9	3	142	43	23	8	42	13	7	3	69	21	11	4	176	54	30	12
N	34	30	25	20	56	51	45	30	142	129	115	80	42	39	35	30	69	63	55	40	176	162	150	120

Allel freq.	(0.95, 0.05)												(0.95, 0.05)²											
s	0.2				0.5				0.8				0.2				0.5				0.8			
m	1	3	5	10	1	3	5	10	1	3	5	10	1	3	5	10	1	3	5	10	1	3	5	10
n	72	23	13	6	115	37	21	9	292	93	53	24	88	28	16	7	142	46	26	12	359	116	67	30
N	72	69	65	60	115	111	105	90	292	279	265	240	88	84	80	70	142	138	130	120	359	348	335	300

Allel freq.	(0.40, 0.30, 0.30)												(0.50, 0.45, 0.05)											
s	0.2				0.5				0.8				0.2				0.5				0.8			
m	1	3	5	10	1	3	5	10	1	3	5	10	1	3	5	10	1	3	5	10	1	3	5	10
n	12	3	2	1	20	5	2	1	54	13	5	1	72	23	13	6	115	37	21	9	292	93	53	24
N	12	9	10	10	20	15	10	10	54	39	25	10	72	69	65	60	115	111	105	90	292	279	265	240

Allel freq.	(0.90, 0.05, 0.05)												(0.90, 0.05, 0.05) (0.40, 0.30, 0.30)											
s	0.2				0.5				0.8				0.2				0.5				0.8			
m	1	3	5	10	1	3	5	10	1	3	5	10	1	3	5	10	1	3	5	10	1	3	5	10
n	88	28	16	7	142	46	26	12	360	116	67	30	88	28	16	7	142	46	26	12	360	116	67	30
N	88	84	80	70	142	138	130	120	360	348	335	300	88	84	80	70	142	138	130	120	360	348	335	300

Allel freq.	(0.90, 0.05, 0.05)²												(0.90, 0.05, 0.05)⁴											
s	0.2				0.5				0.8				0.2				0.5				0.8			
m	1	3	5	10	1	3	5	10	1	3	5	10	1	3	5	10	1	3	5	10	1	3	5	10
n	105	34	20	9	169	55	32	14	428	138	81	37	122	39	23	11	197	64	37	17	496	161	94	44
N	105	102	100	90	169	165	160	140	428	414	405	370	122	117	115	110	197	192	186	170	486	483	470	440

* The description (0.90, 0.05, 0.05)², for instance, shows that three alleles at each of two independent loci are segregating with frequencies 0.90, 0.05 and 0.05.

Table 3. Genotypic frequencies in target populations with 18 different combinations of allelic frequencies and selfing rates

Allelic frequency	Selfing rate	Genotypic frequency					
(p_1, p_2) or (p_1, p_2, p_3)	s	G_1	G_2	G_3	G_4	G_5	G_6
(0.70, 0.30)	0.2	0.5133	0.3733	0.1133			
	0.5	0.5600	0.2800	0.1600			
	0.8	0.6300	0.1400	0.2300			
(0.80, 0.20)	0.2	0.6578	0.2844	0.0578			
	0.5	0.6933	0.2133	0.0933			
	0.8	0.7467	0.1067	0.1467			
(0.90, 0.10)	0.2	0.8200	0.1600	0.0200			
	0.5	0.8400	0.1200	0.0400			
	0.8	0.8700	0.0600	0.0700			
(0.95, 0.05)	0.2	0.9078	0.0844	0.0078			
	0.5	0.9183	0.0633	0.0183			
	0.8	0.9342	0.0317	0.0342			
(0.40, 0.30, 0.30)	0.2	0.1867	0.2133	0.2133	0.1133	0.1600	0.1133
	0.5	0.2400	0.1600	0.1600	0.1600	0.1200	0.1600
	0.8	0.3200	0.0800	0.0800	0.2300	0.0600	0.2300
(0.90, 0.05, 0.05)	0.2	0.8200	0.0800	0.0800	0.0078	0.0044	0.0078
	0.5	0.8400	0.0600	0.0600	0.0183	0.0033	0.0183
	0.8	0.8700	0.0300	0.0300	0.0342	0.0017	0.0342

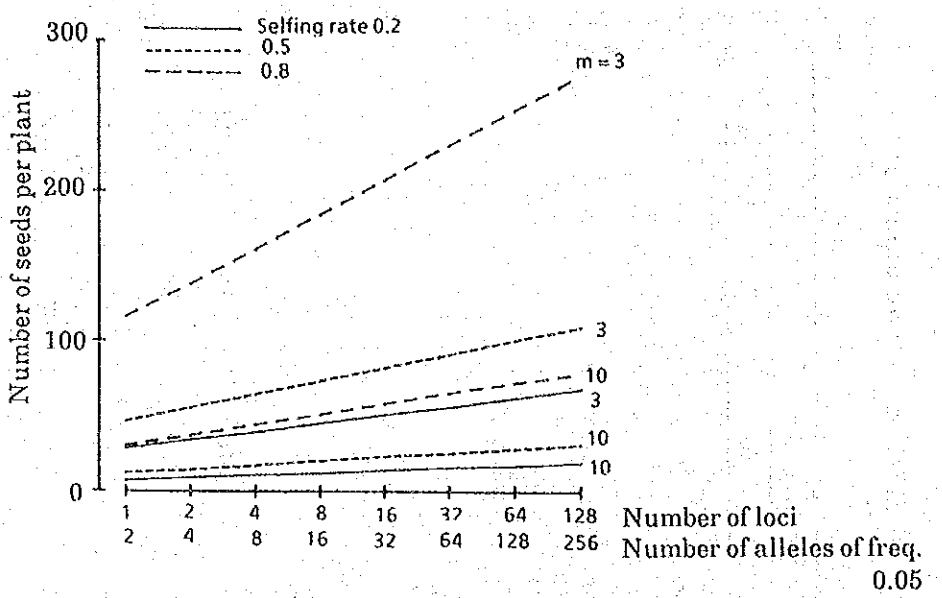


Fig. 3. Number of seeds per plant to be collected in order to give a probability as high as 0.95 of taking up all alleles. Three alleles of frequencies 0.95, 0.05 and 0.05 were assumed for each locus.

b. Sampling for genotypes

The number plants to be sampled to meet P_{IIa} and $P_{IIb}=0.95$ were calculated for similar combinations of allelic frequencies and selfing rates as in the preceding section a. Results are presented in Table 4.

Comparison of Table 4 with 2 reveals that, as intuitively expected, much more plants have to be collected for sampling for genotypes than for alleles. The plant numbers satisfying P_{IIa} and $P_{IIb}=0.95$ largely vary with genetic constitution of target populations, but some similar trends as in the case of sampling for alleles could be pointed out from Table 4. The sample sizes calculated for the cases of a single locus are in the order of (0.70, 0.30)<(0.40, 0.30, 0.30)≒(0.80, 0.20)<(0.90, 0.10)<<(0.95, 0.05)≒(0.50, 0.45, 0.05)<<(0.90, 0.05, 0.05), indicating that the optimum sample size is determined practically not by the total number of different genotypes but by the number of rarest ones. The influence of the number of loci is not very prominent; differences between cases (0.90, 0.10) and $(0.90, 0.10)^2$, between (0.95, 0.05) and $(0.95, 0.05)^2$, and between (0.90, 0.05, 0.05), $(0.90, 0.05, 0.05)^2$, and $(0.90, 0.05, 0.05)^4$ are not so large.

Table 4. Number of plants to be sampled to give a probability (P_{IIa}) as high as 0.95 of collecting all genotypes included in a target population

Allelic frequency	(0.70, 0.30)			(0.80, 0.20)		
Selfing rate	0.2	0.5	0.8	0.2	0.5	0.8
	25	18	21	51	31	29
	(0.90, 0.10)			(0.90, 0.10)[2]		
	0.2	0.5	0.8	0.2	0.5	0.8
	149(149)*	74(74)	56(49)	182	91	66
	(0.95, 0.05)			(0.95, 0.05)[2]		
	0.2	0.5	0.8	0.2	0.5	0.8
	384(384)	162(162)	111(94)	471	199	131
	(0.40, 0.30, 0.30)			(0.50, 0.45, 0.05)		
	0.2	0.5	0.8	0.2	0.5	0.8
	32(25)	29(24)	56(49)	384(384)	172(162)	232(199)
	(0.90, 0.05, 0.05)			(0.90, 0.05, 0.05) - (0.40, 0.30, 0.30)		
	0.2	0.5	0.8	0.2	0.5	0.8
	710(673)	898(898)	1796(1796)	710	898	1796
	(0.90, 0.05, 0.05)[2]			(0.90, 0.05, 0.05)[4]		
	0.2	0.5	0.8	0.2	0.5	0.8
	850	1101	2201	995	1307	2616

* Figures in parentheses represent the plant number to be sampled for taking the rarest genotype (Calculated for $P_{IIb} = 0.95$).

Contrasting to the case of sampling for alleles, the effect of selfing rate on the sample size is not unique. While the sample size for collection of alleles always increases with increase of selfing rate, that for collection of genotypes either increases or decreases depending on the genetic constitution of target populations. In collection of genotypes, the sample size is determined essentially by the frequency of rarest genotypes, which, as seen from Table 3, does not always take a smaller value with a higher rate of selfing.

(3) Discussion

Computations in the preceding section confirm that the number of plants rather than the number of seeds collected per plant primarily determines the success of sampling for alleles (Fig.1). The plant number plays an essential part in keeping the allelic multiplicity of highly selfing populations. In most practical exploration projects, however, the plant number would be limited to some small value due to various limiting factors. In this situation, increasing the number of seeds from single plants is the only practicable alternative. The merit of this approach is expected to be large if the population is not highly selfing ($s<0.8$). A working standard may be derived from computations shown in Table 2 that sampling a few hundred seeds per population, putting emphasis on the plant number rather than seeds per plant if the population is highly selfing, will be sufficient in most if not all circumstances.

The number of plants necessary for sampling for genotypes, on the other hand, is much larger, i.e., mostly of the order of hundreds. It greatly depends on the genetic constitution of target populations, being of the order of thousands when some very rare genotypes are to be collected (Tables 3 and 4).

ALLARD(1970) proposed a sample size of 200 plants×10 seeds per collection site for wild oat populations. This sample size may not always be practiced in actual field expeditions and is estimated to be excessive as far as sampling for alleles is concerned. QUALSET(1975) suggested a sample size as large as 500 plants, which also would be oversufficient for sampling for alleles. His calculation was based on the probability that at least one plant of very rare genotypes (of a frequency less than 0.01) is taken up in the sample. This probability is equivalent to the probability P_{IIb} of this report. As already mentioned, sample size based on this probability may be adequate for some scientific survey, but excessive for ordinary field collection projects.

MARSHALL and BROWN(1975) calculated the number of gametes to be sampled for collecting alleles, and gave a conclusion that a sample of 50-100 plants would be sufficient in most circumstances. The number of seed embryos n was not properly considered in this estimation. HAWKES(1980) seems to have followed this estimation. In his manual for field expedition, he adopted a sample size of 100plants×50seeds for a highly heterogeneous population, and 50plants ×50seeds for fairly homogeneous population. The theoretical ground for the seed

number 50 was not described. According to the calculations based on the criteria of this report, sample size for collection of alleles is estimated to be even smaller.

3) Optimum sample size for multiple target populations

Field collection usually is operated not for a single target population but for a wide area in which there are many collection sites (populations) to be visited. And time and facilities are limited in most collection projects. In this situation, collectors have to make a decision whether to visit many sites spending a relatively small amount of effort to individual sites, or in contrast, to do an intensive collection for a few sites. A criterion to answer this question is presented below.

(1) Mathematical model

Effectiveness of sampling for a target population could be measured by how far the aim of sampling was accomplished by the sampling. As detailed in YONEZAWA(1985), the degree of accomplishment could be defined in terms of any criterion quantity which, as the four probabilities described in the preceding section of this report, was used to measure the outcome of sampling. The degree of accomplishment for a target population was defined so as to take a minimum value (positive and close to 0) when a minimum amount of effort is spent for sampling the population, approaching unity with a larger amount of effort being spent for the sampling. The amount of effort for a target population is proportional to the size of sample.

The effort or resources consist of a number of cost factors such as time, labor and facilities, and will be measured in different ways according to circumstances. It may simply be measured by the amount of one or two factors which most severely limit the scale or coverage of the collection project. Time probably would be the most stringent limiting factor in many field collection projects.

Sampling for a target population should be carried out so that the overall accomplishment of a collection project may be maximized with a total permissive amount of resources. In YONEZAWA(1985), the overall accomplishment, say R, was defined as,

$$R = (\text{Number of populations visited}) \times (\text{Accomplishment for a target population})$$

$$= \begin{pmatrix}\text{Total permissible}\\ \text{amount or resources}\end{pmatrix} \div \begin{pmatrix}\text{Resources spent for}\\ \text{a target population}\end{pmatrix}$$

$$\times \begin{pmatrix}\text{Accomplishment for}\\ \text{a target population}\end{pmatrix} = \frac{E_T}{k(m+c)} \cdot r$$

where,

E_T = total permissible amount of resources for the whole collection project,

k = resources spent per plant in doing collection for a target population,

c = resources (measured in units of k) spent for moving to new sites, transport, facilities etc. which increase proportionally with the number of sites visited,

m = number of plants (sample size) to be drawn from a target population, multiplication $m \times k$ giving the amount of effort spent for a target population,

r = degree of accomplishment for a target population, which is an increasing function of m.

The accomplishment for a target population, r, is defined using the criterion quantity by which the result of sampling is measured. It is formulated as,

$r = I/I_{max}$ ..... for a quantity which increases with increasing size of sample

$I - J/(J_{max} + 1)$ ... for a quantity which decreases with increasing size of sample

where,

I = value of the criterion quantity adopted, which increases as the sample size m gets large.

I_{max} = maximum value of I realized with an infinitely large sample size,

J = value of the criterion quantity adopted, which decreases with increasing size of sample,

J_{max} = maximum value of J realized with minimum sample size, i.e., sampling of only one plant.

By the above definition, the probabilities P_{Ia}, P_{Ib}, P_{IIa} and P_{IIb} can be regarded as the accomplishment r itself, since the probabilities approach unity as the sample size m gets large, i.e., $I_{max}=1$. Besides these probabilities, some other quantities such as expected deviation of genotypic frequencies in sampled plants from true ones in target population (YONEZAWA 1985), expected number of alleles collected from a target population (FRANKEL and SOULÉ 1981), and the probability of some particular alleles being collected (MARSHALL and BROWN 1975), are investigated below. These quantities are presented respectively as,

$$\bar{d} = (g-1)/m$$

$$\bar{u} = u - \sum_{i=1}^{u} (1-p_i)^{2m}$$

$$P_a = 1 - \sum_{i=1}^{u'} (1-p_i)^{2m} + \sum_{\substack{i,j=1 \\ (i<j)}}^{u'} (1-p_i-p_j)^{2m} - \cdots$$

$$- \sum_{\substack{i,j,\cdots,k=1 \\ (i<j<\cdots k)}}^{u'} (1-p_i-p_j-\cdots-p_k)^{2m} + (1-\sum_{j=1}^{u'} p_i)^{2m}$$

where

g = number of different genotypes included in the target population,

m = number of plants sampled,

u = number of all different alleles included in the target population,

u' = number of particular alleles to be collected ($u' < u$),

p_i = frequency of the ith allele at the locus at issue
($p_1 + p_2 + \ldots + p_u = 1$, $p_1 + p_2 + \ldots + p_{u'} \leqq 1$)

By the definition, the accomplishment r for the respective quantities is,

$$r(\bar{d}) = 1 - (g-1)/(gm)$$

$$r(\bar{u}) = 1 - \sum_{i=1}^{u} (1-p_i)^{2m}/u$$

$$r(P_a) = P_a$$

(2) Numerical computations

The sample size m to maximize the overall accomplishment R was computed for some probable combinations of the constituent parameters. Results of the computations are presented in Table 5, where the sample size to give a sufficiently large accomplishment ($r = 0.95$) for individual populations were also shown for reference.

It is seen in this Table that the sample sizes maximizing R are smaller, much smaller when rare alleles or genotypes are involved, than those necessary to satisfy $r = 0.95$. Optimum sample sizes for collection of alleles are remarkably small; sample sizes around or even smaller than 10 are optimum unless very infrequent alleles are to be collected. On the other hand, optimum sample sizes for collecting genotypic variation of target populations, which were calculated based on the quantity P_{IIa}, are larger and more varied with genetic constitution of target populations. It may safely be said, however, that a sample size around fifty plants is sufficient unless very rare genotypes are involved.

Table 5. Sample sizes (m) maximizing the overall efficiency R and achieving a sufficiently large value of the accomplishment r for individual populations

Criterion : condition	Sample size maximizing R			Sample size
	c=1	25	100	achieving r=0.95
$\tilde{d}$: g= 4	2	5	9	15
10	2	6	10	19
50	2	6	11	20
$\tilde{u}$: [u, t] or $(p_1, p_2, \cdots, p_u)$				
= [2, 1.0]	1	3	4	2
[2, 0.2]	1	5	8	6
[4, 1.0]	2	5	7	5
[4, 0.2]	2	5	11	13
[4, 0.03]	1	4	7	80
(0.49, 0.49, 0.01, 0.01)	1	3	29	115
[10, 1.0]	3	10	15	14
[10, 0.2]	3	10	16	20
P_a : $(p_1, p_2, \cdots, p_u)$				
= (0.25, 0.25, 0.25)	4	7	9	8
(0.25, 0.25, 0.25, 0.25)	5	8	10	8
(0.33, 0.33, 0.33)	3	5	7	6
(0.33, 0.33, 0.33, 0.01)	10	43	75	149
(0.49, 0.49, 0.01)	10	43	75	149
(0.49, 0.49, 0.01, 0.01)	64	84	112	183
(0.63, 0.23, 0.09)	5	11	17	16
(0.80, 0.05, 0.05)	14	23	33	36
(0.80, 0.05, 0.05, 0.05)	20	28	37	40
(0.80, 0.05, 0.05, 0.01)	33	49	76	149
P_{Ia} : (p_1, p_2), s=0.2, n=5				
= (0.99, 0.01), s=0.2, n=5	6	23	38	51
(0.95, 0.05), s=0.2, n=5	2	8	12	11
$(0.95, 0.05)^2$, s=0.2, n=5	5	10	14	13
(0.90, 0.10), s=0.2, n=5	2	5	7	5
$(0.90, 0.10)^2$, s=0.2, n=5	3	6	8	7
(0.99, 0.01), s=0.8, n=5	9	39	68	129
(0.95, 0.05), s=0.8, n=5	4	15	23	26
(0.90, 0.10), s=0.8, n=5	3	9	14	13
(0.99, 0.01), s=0.5, n=1	11	47	84	179
(0.95, 0.05), s=0.5, n=1	5	18	29	36
(0.90, 0.10), s=0.5, n=1	3	11	18	18
(0.99, 0.01), s=0.5, n=50	2	9	13	12
(0.95, 0.05), s=0.5, n=50	1	3	4	3
(0.90, 0.10), s=0.5, n=50	1	2	2	2
P_{IIa} : (p_1, p_2), s=0.2				
= (0.95, 0.05), s=0.2	39	74	133	384
$(0.95, 0.05)^2$, s=0.2	162	187	234	471
(0.90, 0.10), s=0.2	20	43	75	149
$(0.90, 0.10)^2$, s=0.2	64	83	112	182
(0.80, 0.20), s=0.2	10	23	37	51
(0.70, 0.30), s=0.2	7	15	23	25
(0.95, 0.05), s=0.5	36	55	82	162
(0.90, 0.10), s=0.5	18	32	49	74
(0.80, 0.20), s=0.5	10	18	27	31
(0.70, 0.30), s=0.5	7	13	18	18
(0.95, 0.05), s=0.8	40	56	76	111
(0.90, 0.10), s=0.8	21	33	45	56
(0.80, 0.20), s=0.8	11	19	27	29
(0.70, 0.30), s=0.8	8	15	21	21

t= Relative frequency of one allele compared to the other u - 1 alleles,
e.g., [4, 1.0]=(0.25, 0.25, 0.25, 0.25), [4, 0.2]=(0.3125, 0.3125, 0.3125, 0.0625)

The sample size maximizing R becomes smaller as the cost parameter c, i.e., the amount of cost necessary for traveling and transport between sampling sites, gets small. For quantities $\hat{d}$ and $\hat{u}$, the optimum sample size is influenced more sensitively by the cost parameter than by the genetic constitution of target populations.

(3) Discussion

Numerical computations in the preceding section showed that, while some fifty plants as a rough estimate have to be drawn for collecting genotypic multiplicity of target populations, a strategy of collecting fewer plants from more populations is efficient for collection of allelic diversity, indicating that unless traveling and transport between sites are very costly, visiting as many sites as possible is more rewarding than doing intensive sampling for a limited number of sites. Some other quantities such as chance fluctuation of allelic frequency in samples and the degree of heterozygosity maintained in the sample (cf.,e.g., CROW 1954) may be used to measure goodness of sampling. Numerical computation by these quantities, although not presented in this report, led to the same conclusion.

Sample size for multiple target populations has initially been discussed by OKA(1975) based on the probability (G) that at least one copy of a particular allele aimed is collected, which was formulated as,

$$G = 1 - \{(1-P) + P(1-p)^n\}^N$$

where

P = frequency (among all populations in the whole target area) of the populations which contain plant genotypes carrying the aimed allele,

p = frequency of genotypes carrying the aimed allele in each of such populations,

N = number of populations sampled,

n = number of plants sampled from a population.

Assuming that alleles at different loci are independently inherited and randomly distributed over all target populations, the above quantity G may be interpreted as the expected fraction of loci at each of which aimed alleles have been taken up.

Application of the probability G is considered to be rather limited. Collecting particular alleles from any one population in the target area would not be the most important subject of most field expedition projects. Taking an unbiased, though small, sample from each of the target populations to infer some genecological features would mostly be the main issue. The probability G includes no parameter or quantity which measures, and therefore imposes any condition on, the result of sampling for individual populations. The criterion R introduced in this report is defined using the accomplishment for individual populations, r, and is compatible with any quantities which have been proposed to measure the goodness of sampling for individual populations.

It should be noted in this context that, by OKA(1975) himself, the parameters P and p and consequently G were not interpreted as above. In his definition, the parameters P, p and G respectively are the fraction of genetic variation represented by a population, that represented by a plant within a population, and that represented by the total N·n plants. Following this definition, he actually estimated P and p by the ratio of within population genetic variance to the overall (within plus between population) genetic variance, and the ratio of within progeny genetic variance to the total (within and between progeny lines) genetic variance of a population, respectively. This interpretation of the parameters is invalid. A detailed discussion would be necessary to confirm this strictly. In this report, it would be suffice to show some contradictions coming out by OKA's definition.

Now suppose that $P=1$ and $p=0$, which in OKA's definition correspond to the situation where there is no genetic difference among populations (i.e., every population is genetically identical and contains all genetic variation available) and no genetic variation within progeny lines of plants (i.e., all plants within populations are homozygous). In this situation, the quantity G becomes zero irrespectively of the values of N and n, meaning that no genetic variation can be collected with any scale of sampling. This obviously is invalid. G is always zero when $P=0$, i.e., when no genetic variation exists within populations (all plants within populations have an identical genotypes whether homozygous or heterozygous). This means that no genetic variation can be taken up with any

large scale of collection when target populations are genetically homogeneous. This also is contradictory.

In spite of the above difficulty, OKA's criterion, when allocation of effort is taken into account (MARSHALL and BROWN 1983), makes essentially the same conclusion as the one derived in this report for collection of alleles; the number of populations rather than the sample size for a population determines the overall efficiency of a collection project.

4) Summary

Sample size for collecting germplasms (alleles and genotypes) from natural plant populations was discussed. Genetic potential of a target population could be maintained, or somehow recovered when necessary, if all important alleles in the target population have been collected. So, it is allelic rather than genotypic multiplicity that should be aimed at in most field collection projects.

Sufficient size of sample for collecting allelic multiplicity of a target population was investigated in relation to the numbers of plants and seeds per plant to be sampled. It was derived that in predominantly selfing populations goodness of sampling is primarily determined by the number of plants rather than seeds per plant, since the genotypes of seed embryos produced on a highly selfing plant are highly homozygous and homogeneous. The number of plants, however, does not need to be large. The drawback of a shortage in the plant number can be avoided by collecting sufficiently many seeds from each plant. Computations for some probable situations led to the conclusion that a few plants per population may be enough if the plants bear a few hundred seeds each and are not highly selfing. This sample size is much smaller and more practicable than those proposed previously.

Not alleles but plant genotypes in a target population have to be unbiasedly sampled for a research survey project to infer frequencies and geographic distribution of plant genotypes. A much larger sample size is required in this situation than in case of collection for allelic multiplicity mentioned above. Sufficient number of plants for collection of genotypes largely depends on the genetic constitution of target populations. A sample size of the

order of hundreds or more is needed when some very rare genotypes (of a frequency as small as or smaller than one percent) are involved.

Time, labor and facilities are limited in most if not all field expedition projects. Collectors therefore have to decide whether to visit many populations (sites) with a relatively small number of plants being collected from single populations, or in contrast, to do an intensive collection for a few populations. In this situation, a strategy of sampling fewer plants from more populations (sites) may generally be recommended for collection of allelic multiplicity, while a sample size of around fifty plants as a rough estimate seems to be appropriate to draw genotypic multiplicity of target populations.

5) Literature cited

ALLARD, R.W., S.K.JAIN and P.L.WORKMAN 1968. The genetics of inbreeding populations. Adv.in Genetics 14:55-131.

ALLARD, R.W.1970. Population structure and sampling methods. In 'Genetic Resources in Plants—Their Exploitation and Conservation' ed. by O.H.FRANKEL and E.BENNET, Blackwell, Oxford and Edinburgh.97-107.

BROWN, A.H.D.1978. Isozymes, plant population genetic structure and genetic conservation. Theor.Appl.Genet. 52:145-157.

CROW, J.F.1954. Breeding structure of populations. II.Effective population number. In 'Statistics and Mathematics in Biology' ed. by O.KEMPTHORNE, T.A.BANCROFT, J.W.GOWEN and J.L.LUSH, Iowa State Univ.Press, Ames. 543-556.

ECKHOLM, E.P.1976. Losing Ground—Environmental Stress and World Food Prospective (translated into Japanese by H.ISHI and K.MIZUNO). Aoki Shobo, Tokyo.

FRANKEL,O.H.and M.E.SOULÉ 1981. Conservation and Evolution. Cambridge Univ.Press, Cambridge.

HAWKES, J.G.1980. Crop Genetic Resources Field Collection Manual. Dept.of Plant Biology, Univ.of Birmingham, England.

HAWKES, J.G.1983. The diversity of Crop Plants. Harvard Univ. Press, Cambridge, Massachusetts.

JAIN, S.K.and D.R.MARSHALL 1967. Population studies in predominantly self pollinating species. X.Variation in natural populations of *Avena fatua* and *A. barbata*. Amer.Nat. 101:19-33.

MARSHALL,D.R. and A.H.D.BROWN 1975. Optimum sampling strategies in genetic conservation. In 'Crop Genetic Resources for Today and Tomorrow' ed.by O.H.FRANKEL and J.G.HAWKES, Cambridge Univ.Press, Cambridge, 53-80.

MARSHALL, D.R. and A.H.D.BROWN 1983. Theory of forage plant collection. In 'Genetic Resources of Forage Plants' ed.by J.G.MCLVOR and R.A.BRAY, CSIRO, Melbourne, 135-148.

MORISHIMA,H.,Y.SHIMAMOTO, Y.SANO and Y.I.SATO 1984. Observation on wild and cultivated rices in Thailand for ecological-genetic study - Report of study-tour in 1983. Contribution from National Institute of Genetics No.1621, Mishima.

OKA,H.I.1975. Consideration on the population size necessary for conservation of crop germplasms. In 'JIBP Synthesis Vol.5' ed. by T.MATSUO, Tokyo Univ. Press, Tokyo, 57-63.

OKA, H.I. and H.MORISHIMA 1967. Variations in the breeding systems of a wild rice, *Oryza perennis*. Evolution 21:249-258.

QUALSET,C.O.1975. Sampling germplasms in a center of diversity: an example of disease resistance in Ethiopian barley. In 'Crop Genetic Resources for Today and Tomorrow' ed.by O.H.FRANKEL and J.G.HAWKES, Cambridge Univ. Press, Cambridge, 81-96.

RAMADE,F.1981. Ecology of Natural Resources (translated into English by W.J.DUFFIN), John Wiley & Sons, New York.

VASEK,F.C. and J.HARDING 1976. Outcrossing in natural populations. V.Analysis of outcrossing, inbreeding, and selection in *Clarkia exilis* and *Clarkia temblorіensis*. Evolution 30:403-411.

YONEZAWA,K.1985. A definition of the optimal allocation of effort in conservation of plant genetic resources - With application to sample size determination for field collection. Euphytica 34:345-354.

I – 3. Manual for Exploration and Collection of Plant Genetic Resources

by

Tsukasa NAGAMINE and Masahiro NAKAGAHRA

National Institute of Agrobiological Resources

CONTENTS

For Emergency Use

In case of emergency, make an adequate memorandum according to the form described below, and contact with appropriate organizations.

'Do Not Get Upset'

Emergency contact form

1. Information to be transferred
   1) Sender and memorandum number
   2) When (Date, time of accident)
   3) Who (Name of all members were involved in the accident)
   4) What, How (Describe the result of accident briefly)
2. Present condition
3. Provisional measures to be taken
4. Request to head office (If necessary, itemize concretely)
5. Contact methods to the mission (Address, telephone, telex)
6. Next contact (If any, provisional time)

Contact points

1. Head office (Facsimile, telex, telegram, telephone No.,address)
2. Second contact point; Curator (do.) - - - - - - - Inform to 1 (Head office).
3. Third contact point (do.) - - - - - - - Inform to 1 and 2.
4. Members' home address and telephone No.

Members

Name (Roman letters), occupation

1) Introduction

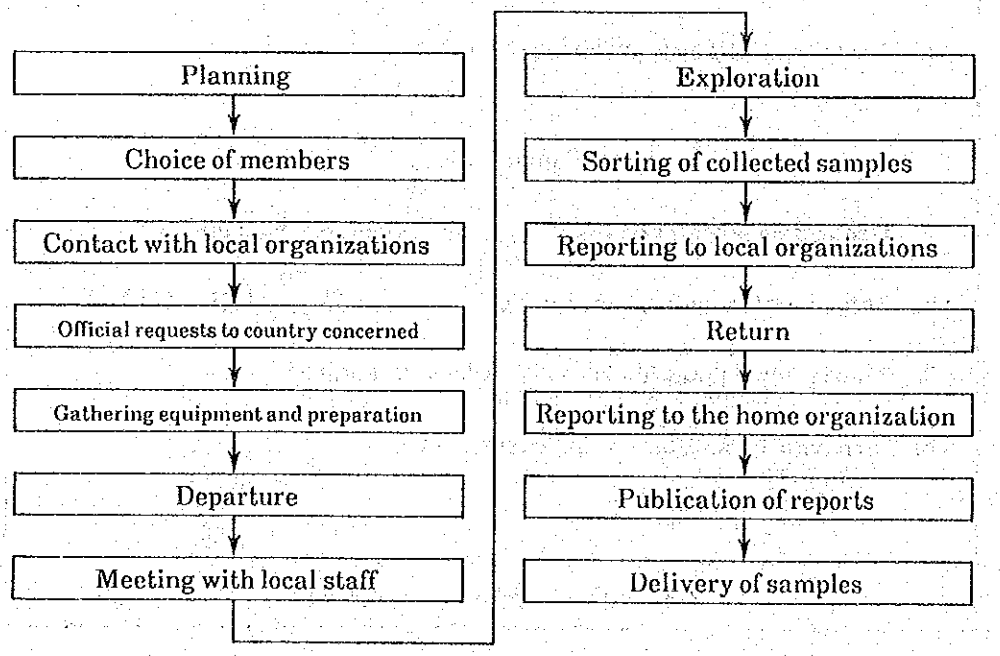
The importance of explorations for plant genetic resources is now widely recognized. The first step in an exploration is to meet the farmers in order to collect information from them and the final stage is to return with the 'Treasure' in the form of improved varieties.

Practice, methods and techniques of exploration and research include extremely intricate elements, and these forms are various. Accordingly, during an exploration, a researcher should be prepared to act flexibly with his/her plans in case any unexpected changes in the situation.

There are several kinds of plant germplasm; cultivars, those relative species and wild species. The size of a collection varies according to the breeding system of the plant. Here, we will mention several topics which are common to all plant explorations.

2) Flow chart of exploration and collection

When one is asked to explore certain plant germplasm, one should make preparation according to the flow chart described below.



3) Strategy

A strategy is needed to determine what kinds of targeted genetic resources are to be collected, who is to collect them, where, how, when, and how long the exploration will take.

(1) What? (Targeted genetic resources)

a. Purpose of exploration

Make clear what the targeted species or crops are. An obscure purpose of exploration leads to limited results. Until several decades ago, this kind of exploration was called a plant expedition, and meant an extensive botanical collection. Recent exploration of plant genetic resources has a clear purpose which is the collection of materials for improving crops. Plant breeders select certain variations and improve crops, on the other hand, plant explorers collect and enlarge the genetic variabilities of crops.

Two kinds of exploration can be considered. a) Single crop exploration b) Multi-crop exploration. The collected samples of the former have a high reliability because the collection has been carried out intensively. The latter involves the collection of a lot of species and its success depends upon the extensive and intensive knowledge of each mission member.

b. Target and scale of collections

The effective collection of genetic variabilities has to deal with characters which are difficult to evaluate at collection sites but which have several potential values. Consequently, the more targeted crop species is and too more samples collected, the better. However the scale of a collection is finally determined by the manageable quantity at the gene bank and the restrictions imposed by the plant quarantine law.

(2) Where? (Choice of area of exploration)

The most effective area of exploration should be determined on the bases of the knowledge accumulated up to now and on that obtained in the preliminary survey. A primary concern of researchers is the distribution and variability of

plant species, however, they must consider the social circumstances in the area, race and religion.

(3) Choice of route and preparation of maps

The planning of the route of exploration requires detailed maps, i.e. those on a scale of 1 to 500,000. The scale of the map required will also vary depending upon one's means of transportation, which is generally either by car or on foot. It is essential to have the maps beforehand. However, it is comparatively difficult to obtain the maps of foreign countries. Therefore we frequently purchase or obtain detailed maps after entering a targeted local area. Even detailed road maps or guide maps for sightseeing can prove to be very useful. At any rate, one must choose the route and make a tentative schedule based on such unsatisfactory maps. During the planning, it is very easy to draw up an excessive schedule, and so one should always remember to make a reasonable realistic schedule.

(4) When? (Suitable season of exploration)

To make a schedule of exploration at an appropriate suitable season is extremely difficult.

a. Season

Exploration and collection of cultivated germplasm largely means visits to farmers. The maturing stage of the targeted germplasm is the most favorable season not only for the collection of seeds but also for obtaining necessary plant characters. As the farmers are working in their fields, one can often obtain from them important information about crops and cultivation. Since seeds of wild species are easily shattered after maturing, collection should be done before the maturing stage. Consequently, the suitable stages of both wild and cultivated species are different from each other, and researchers must plan the itinerary bearing in mind which are the suitable stages for each targeted species or crop.

b. Length of exploration

The length of an exploration depends upon the extent of distribution of germplasm, the extent of the distributed area, transportation means, and

especially the extent of funds. Normally, the length is between two and five months.

When the length of exploration is short, the researchers will tend to visit institutions or farm house and the fields neighboring the town. To avoid unsuccessful collections, one should allocate a few days for preparation for departure, meeting with local staff, sorting out of collected samples and making reports.

4) Preparation before departure

(1) Contact with local organization

Initially contact should be made with an official organization in the area of exploration. This contact should be made with the responsible organization or laboratory, and the possibility of carrying out an exploration in the targeted area should be investigated.

Additionally, if necessary, a final official letter from the director of institute should be provided. The responsible person in the mission should contact the curator (a responsible researcher of each crop) and make necessary preparations. When visiting a new area, one may feel uneasy even though one has got sufficient information. It may not be possible to carry out an effective and efficient exploration on the first visit, but try one's best.

(2) Acquisition of minimum information

Survey the local conditions through discussion with an experienced person and surveying the literature. Sort out any kind of available information, investigate the distribution of the crop species and the extent of their diversity. Also be aware of any influence on the local circumstances played by social, racial and religious customs, and note the different custom of country generally. As there is often little literature available about conditions in most of targeted areas of exploration, one may have to start with the minimum amount of information. It often happens that the office hours and national holidays are considerably different from a country to another. If one enters the area without such basic information, one is generally obliged to change one's itinerary later.

Therefore one should find out at least the minimum information about a targeted area.

(3) Pre-studying of existing germplasm

Before departure, one should investigate the existing collections of seeds or herbaria. If one is unfamiliar with the targeted species, one should visit the gene banks or field gene banks and observe the seed samples of living material. Then one can define the uncollected part of the targeted species and understand the important points about the forthcoming exploration.

5) Travel planning

(1) Membership

There has been considerable discussion about the membership of any group for an exploration. FRANKEL and HAWKES (1975) suggested that only an explorer and a translator is sufficient. On the other hand, it has been argued that an exploration mission should consist of several members including taxonomists, physiologists and plant pathologists[15]. In our experience, most exploration missions consist of two scientists who have the same speciality and one person who is well informed on the local conditions. In the case of multi-germplasm exploration when the group consists of several researchers who have different research fields, the group sometimes explores separately which is not very efficient. The high efficiency of collection in a local area is generally accomplished by the cooperative work of several researchers.

The membership of a group should of course include staff from the local institute as this is the most effective means of finding excellent local staff for the exploration. Additionally, at least one expert on the targeted crops should be included in a mission, and it is desirable that some young researchers who will later become one of the active people in future explorations, are included. As youthfulness is of some importance in exploration activities, we should train the next generation as soon as possible.

(2) Preparatory meetings and assignment of roles

It is necessary to call the preparatory meetings in some detail. To ensure the success of an exploration mission, hold the preparatory meetings several

times and decide what the role is of each person. This means that at the first meeting the role of each person should be assigned; i.e. team leader, those responsible for liaison, finance and procurement and so on and each responsibility made clear. A timetable for preparatory work in a home institute should be made. This is the key for success.

(3) Purchase of basic equipment

The person in charge of equipment should prepare the equipment according to the type of mission. One should check it carefully before departure, use it beforehand, and sort out any shortages. Make a point of finding out what each member will be bringing regarding their own. A little as possible of the members own living equipment should be brought, as most of these should be purchased locally.

(4) Contact with plant quarantine office

Contact the plant quarantine office in advance to promote a smooth introduction procedure, if necessary. Submit a note of the quantity of inspecting plants and presence/absence of prohibited plants to be imported. Then, coordinate with quarantine officers about the quantity and so on. Submit a planning form for the introduction of plant genetic resources to the plant quarantine office. If one is going to return by air, send the copy to the branch office of plant quarantine at the airport.

6) Equipment

There are certain variations in equipment required for exploration activities depending on the purposes and methods of collection. Table 1 shows a list of equipment. This is one example. In brief, choose small and light equipment, it is better to bring familiar equipment, and it is unnecessary to bring a lot of special equipment which one does not normally use at one's institute. Unusual equipment might embarrass local farmers. Prepare equipment that is possible to carry in vehicles. When an airplane is used, make it as small as possible.

(1) Collecting equipment

Small size seed bags, scissors, knives, staplers, collection form books, maps, diaries, several cloth bags, net bags and so on. Prepare these according to the crops and species that one wants to collect.

(2) Measuring equipment

Several measures, seed cleaners, compasses, altimeters, hardiness-meters, pH meters, cameras, films and so on. Discuss with the mission members to determine who possesses these equipment.

(3) Lodging equipment

There will be differences in equipment according to whether the mission will move from site to site by car and stay in accommodation or travel on foot and stay in tents. Generally speaking, one should stay in a single room at a lodging or use a personal tent in a long exploration, because one needs some private time and space for refreshment.

(4) Transportation

Don't expect the whole area where an exploration will be carried out to have good transportation conditions. Use all transportation means available and ensure that transfer can be made. SIMPSON *et al.*(1983) used several kinds of transportation; airplanes, charter planes, buses, passenger trains, freight trains, cars, trucks, jeeps, motorcycles, ships, lighters, dugouts, house wagons, cable cars and also had to walk a lot during their peanut exploration from 1976 to 1983. In our experience we have used all the equipment described above except dugouts. Transportation conditions are similar everywhere. In any case, transportation is one of the important factors leading to the success of exploration. HARLAN[14] suggested an ideal scheme in which a mission should initially use a helicopter as the main means of transportation, and decide on base villages, then move from there by tough vehicles to the locations that can be driven through and finally walk within the area. Exploration activities go into unexplored land. Availability of transportation in a local area is very rare and is often less than ideal. Make as much contact as possible with local organizations. If one is still anxious about it, believe in one's 'good luck'.

(5) Clothing

It depends on season and length of stay. Prepare, at first, tough shoes. Light mountain-climbing shoes, for instance, are very convenient and acceptable not only for collection activities but also for visiting institutes. One pair of these shoes will be sufficient for a whole range of activities.

Bring a small number of durable and washable jackets, shirts and trousers, and underwear.

(6) Medicine

Prepare several kinds of medicine to cope with emergencies. Don't be afraid to take a large amount unless the weight of medicine becomes too heavy, as local villagers will occasionally ask for urgent medical treatment, and so a large quantity of ointments and eye lotions would please them.

Table 1 List of equipment for exploration

Item	Articles	Notes
Collecting	Sample bags (cloth, net, polyethylene and others), Collection forms and Summarizing tables, Diary, Maps, Road-maps, Staplers, Knives, Scissors, Rucksacks, Field aprons	It depends on crops/species. Reserve pins!
Measuring	Altimeters (5,000m), Compasses, Cameras, Films, Several measures, Seed cleaners, Sterilizers, Desiccants, Portable warmers, Portable coolants, Containers, Range finders, Photometers, Hardiness-meters, Magnifying glasses, pH meters, Goods for making herbaria	if, necessary
Lodgings	Tents (mats, flies), Sleeping bags, Lamps, Candles, Water bottles, Portable cooking stove, Cooking instruments	can be purchased in local area
Transportation	Packing materials, Transporting box, Vehicles (4WD)	
Clothing	Jackets, Shirts, Trousers, Underwear	devise by oneself!
Medicine	A medical set for emergencies, Antibiotics, One's own medicine	for chronic disease

Note: As these articles will depend upon the type of exploration. Investigate fully beforehand!

7) Methods of exploration and collection

(1) Kinds of collection sites

a. Places with wild species growing spontaneously

When targeted species are genuinely wild, they grow in areas away from human habitation. These are hilly places, valley regions, riverbeds, forests and surrounding areas, and seashores. Although the flora of these regions have occasionally been changed by pasturage, check beforehand whether these areas are pasturelands or not. In this sense, since national parks and national reserves are non-pasturelands, artificial disturbance is small, researchers will have a good chance of obtaining unexpected populations.

b. Farmers' fields

When one wants to collect several cultivars and related wild species, one should mainly visit farmers' fields. Here, an activity begins with an interview with farmers rather than with the collection. Behave politely and talk to them sincerely and tactfully. Explain the intention of one's visit carefully so that they can understand the real purpose. After that, ask them to show one their fields, and start the survey of crops and other plants. At the edge of fields, the threshing places near villages and seed houses, researchers can collect the seeds of crops which they missed at the appropriate harvesting time. In such cases, as these samples frequently contain seeds of different varieties, one should ask for information about seeds concerned, and try to obtain accurate details about collected seeds.

c. Markets and open markets

If genuinely domestic cultivars are marketed, the mission can increase the number of collections by buying these varieties. In seasonally opened outdoor markets and free markets, farmers bring their harvested crops. These market collections are very valuable. The interview about seed origin, however, should be done carefully. One has to accept that occasionally one will receive wrong information and/or seeds.

d. Institutions

When one visits certain crop research institutes, they can sometimes share the materials which already conserved in those gene banks. These collections seem very effective because one can obtain a lot of collections at once. However, in the view of genetic resources exploration number of accessions do not increase as a whole. Hence this kind of 'present' is not 'collection' or 'exploration'. Some of exploration missions even now carry out this type of collection, and host institute find fault in such behavior.

(2) Decision of collection sites

a. General

If there exists little geographical variation collection should be carried out over a certain distance. Generally speaking, collection should be carried out at the distance of 50km in the case of less variable regions.

b. Topography

Considerable differences of growth environment for plants exist between the closely locating two villages separated by very different geographical features. In these villages wide variation of genetic resources can be expected. Consider the topographical features carefully and start one's collection. Since the same methods can be applied to the adjacent areas having large difference of altitude, carry out collections with specified distance.

c. Special areas

If an area is inhabited by different tribes, considerable variations of agriculture and dietary custom will exist in a very small area. Therefore, we can expect several collections that are constituted with not only the genetic variability but also the diverse cultural differences. We had collected these variants in South-east Asia, India and Nepal. Collect the germplasm carefully with bearing in mind this variability.

d. Sampling methods

The standard of sampling methods varies according to the targeted germplasm; seeds, fruits, scions and bulbs [1,2]. The ideal method is to increase

the efficiency of collecting without causing distortion to the diversity in the population, however the ideal method and practical one were different from each other. Here, we concentrate on seed crops mainly.

a) Types of sampling methods based on the objective of the collection

The types of sampling methods differ according to the objectives of the collection.

(a) Collection for detecting the influence of geographical features and special areas on a population.

(b) Collection for investigating the genetic structure within and between populations in the explored sites.

(c) Collection of small amounts of seed on the assumption that collected seeds will be regenerated at first and not used as genetic resources directly.

(d) Conservation of the collected seeds into the gene bank directly for subsequent utilization.

(e) Collection for direct regeneration.

(f) Collection for obtaining particular individuals for breeding purposes.

The above-mentioned types can be summarized into two categories; i. conservation of all collected samples as genetic resources with a view of using materials directly, ii. regeneration of the seeds of collected samples after bringing back, which could then be conserved and utilized as genetic resources. From the view point of safe conservation the variations of the genetic resources, the first category should be the standard. However, this method can not be adopted for the species and crops the import of which is prohibited and/or those designated to be kept in isolation after introduction.

b) Sample size

Sample size should be as large as possible. Sample size is the eternal question for the plant explorers although some proposals for field collections were made elsewhere [14,17,18]. The standard which has been derived from

experience is 2,500 seeds or more of 50-100 individuals for the small grain plants should be collected from one site[1]. It generally takes much longer time than expected to collect in fields, consequently the size of a sample is dependent on time available. In our experience of rice collection, samples of seeds of 100-150 plants are standard.

c) Non-selective sampling

If one wants to collect the germplasm without missing any variations in each site, then non-selective sampling is recommended. Seed collections should be made by sampling a small amount of seeds from many divided plots in one site, not by sampling a large amount of seeds from one particular plot. Theoretical study has shown that there is a high probability that non-selective sampling will cover the diversity of alleles in sites and it is consequently rare that valuable genes are lost [14,16,17,18]. Practically, sample 50 seeds from every one plant, select the plants randomly with the interval of 10-20 steps, repeat this procedure until accumulating 50-100 individuals in total [1]. In the case of plants such as buckwheat which yield only several seeds on one plant, regard some adjacent groups of individuals as one plant, and then collect the prescribed amount of seeds. In practice, non-selective sampling in fields is both labor and time intensive, however seed collected by this method is extremely reliable and is of a high value in terms of genetic resources. Non-selective sampling can easily be adopted for the collection of wild species, however in the collection of cultivars there are many problems that are encountered may be that of small land area where insufficient amount of seeds can collected by non-selective sampling, or, that the farmers cultivate mixed varieties intentionally. Before beginning any collection, it is necessary to obtain permission for collecting samples from the farmers, who own the fields. In our experience, the amount of seeds of vegetables stored by the farmers is comparatively small, in many cases the farmers may maintain only dozens of seeds for each crop. Accordingly, the sufficient amount of seeds can not be easily collected in such cases.

d) Selective sampling

This method is to collect some particular materials by selecting the individuals with considering the significance of information from plants. Selective sampling depends upon one's experience. Samples collected selectively are not recognized to be one of representative of diversity in sites. Normally selective sampling should be carried out with non-selective sampling in practice.

In all cases, observe the plants in fields and ask the farmers sufficiently, and judge that the variation is mixed varieties or population which possesses original variability. The above rules vary according to the kinds of sampling organs; seeds, fruits, scions and bulbs. However, the two methods of utilization, namely, (a) direct utilization of collected samples as genetic resources, (b) subsequent utilization after regeneration of collected samples should be obviously distinguished.

Collection of cultivars needs permission from the farmers. The amount of seeds and sampling methods are inevitable to be according to the understanding and allowance of the farmers.

8) Guideline for exploration activities

As the date of departure for exploration become imminent. The guidelines for activity of each day which are summarized below should be followed.

(1) Working hours

In all types of exploration aim to commence activities at sunrise. Have an early lunch and finish traveling 2-3 hours before sunset. It is also best to prepare a packed lunch beforehand. Such guidelines allow for a more efficient and effective use of limited time.

(2) Range of activity

When traveling by car it is possible to cover distances of 100-150 km per day on bad conditioned roads and 200 km on well maintained roads. Whereas on foot, 10 km is standard. When one starts from base-camp, take different routes on going and returning. Measurement of distance traveled is estimated by a range finder when using car and by a pedometer or a map when traveling on foot. As the distribution of species along a well maintained road is often disturbed, it is best to get out of the car and walk into the vegetation. It is also best to ensure beforehand that the range finder is accurate.

(3) Contents of working at lodging

After arriving at the camping sites or lodgings, the following jobs should be carried out immediately.

a. Treatments of collected samples

Sterilization and drying of samples should be done immediately. Where necessary remove the seeds from fruits, and clean the surface of any seeds or bulbs. Later one has no time to manage above-mentioned treatments. Furthermore, seeds also should be cleaned for the sake of plant quarantine inspection and for future utilization.

b. Making herbaria

Make some herbaria, which will be necessary for taxonomical study in the future, from the plants collected on these occasions.

c. Filling in collection forms

The records of collected samples should be completed by the end of the day. The collection forms and summarizing tables should be filled in completely at the lodging.

d. Sorting and packing of collected samples

Sort the collected samples clearly day by day. It is recommended to send back a large amounts of samples or herbaria at regular intervals of a few days, if possible. Therefore, sorting and packing of collected samples is vital.

e. Keeping a diary

Before going to bed, keep one's diary. Record the activity of each day, the map reference of collecting site and collection forms, the amount of distance traveled, any notable observation and so forth.

9) Documentation into collection forms

If detailed records of collected samples is lacking, the value of precious genetic resources will be reduced by half. After the completing an exploration,

the information obtained at collecting site will be forgotten. Therefore continue to record every day 'earnestly', and complete the filling of collection forms at the same time at the end of each day exploration activity.

(1) Collection forms

Since the necessary contents are different from the species of collected samples, it is best to prepare one's forms to cope with the species one collects. The 'General form' proposed by IBPGR is a good example for an unspecific genetic resource (Fig.1). If one wants to collect a particular species/genus, make a new form that is similar in style to that of the 'General form'. For the simple handling of forms, B6 size of sheet is recommended. 50 or 100 sheets of forms to one pad and stamp the sequential numbers on them. Every pad should be covered with a cardboard. Fig.2 shows an example of a form for cereal crops. Furthermore, it is difficult and time-consuming to investigate and record information simultaneously. Therefore such work should always be done in pairs.

(2) Style of summarizing table

Inadequate records of the collected samples will half the value of the samples as a genetic resources. All the information about samples should be recorded before the leaving of the exploring area. The information should then be summarized on a table and reports made to the organizations concerned in the area where the samples have been collected. Details of the exploration should be published promptly. Recording all the necessary information may be time-consuming but it is best to do so whilst all the details are fresh in one's mind. The details which one should fill in the summarizing table will later be used as the basis of the collection lists for the exploration report. Fig.3 shows an example of a summarizing table.

IBPGR COLLECTION FORM (GENERAL)

GENUS:_____
SPECIES:_____
SUBSPECIES:_____

COLLECTOR'S NUMBER:_____
COLLECTING INSTITUTE:_____
*DATE OF COLLECTION:*_____

COUNTRY OF COLLECTION:_____
*PROVINCE/STATE:*_____
LOCATION OF COLLECTION SITE
nearest town/villate:_____
distance(in km):_____
name of village:_____

*LATITUDE OF SITE:*_____ N S
LONGITUDE OF SITE:_____ E W
ALTITUTDE OF SITE:_____ (m)

COLLECTION SOURCE(circle one)

1 wild	5 village market
2 *farmland*	6 *commercial market*
3 farmstore	7 institute
4 backyard	8 others(specify)

STATUS OF SAMPLE(circle one)

1 wild	4 landrace
2 *weedy*	5 cultivar(bred)
3 breeder's line	6 others(specify)

LOCAL NAME:_____

NUMBER OF PLANTS SAMPLED:_____
PHOTOGRAPH (circle one): yes no
photo number:_____
TYPE OF SAMPLE(circle one)
1 vegetative 2 seed 3 both
HERBARIUM SAMPLE(circle one): yes no
QUANTITY OF MATERIAL (number of seeds or plants/sample):_____

CULTURAL PRACTICES(circle one):

shifting	yes	no
irrigated	yes	no
transplanted	yes	no
terraced	yes	no

*SOWING MONTH:*_____
HARVEST MONTH:_____
USAGE(specify):_____

DISEASE AND PEST(specify):_____
_____

TOPOGRAPHY(circle one)Associatea wild and *weedy species and crops* (specify):

1 *swamp*	_____
2 flood plain	_____
3 plain level	_____
4 *undulating*	
5 hilly	
6 mountainous	
7 *other (specify)*	

SITE(circle one)

1 level
2 slope
3 summit
4 depression

STONINESS(circle one)

1 none
2 low
3 medium
4 rocky

SOIL TEXTURE(circle one)

1 sand
2 loam
3 clay
4 silt
5 highly organic

DRAINAGE(circle one)

1 poor
2 moderate
3 good
4 excessive

OTHER OBSERVATIONS:_____
_____
_____
_____

NIAR, MAFF COL.NO.	NIAR, MAFF COL.NO.	NIAR, MAFF COL.No.	NIAR, MAFF COL.NO.

Fig. 1 An example of collection form i.e. 'General form' of IBPGR

NIAR, MAFF

COLLECTION No.: DATE: / PERSON:
GENEUS and SPECIES:

LOCAL NAME(CULTIVAR NAME):
SAMPLE: seed vegetative / indiv. popul.( ) / quantity g/ herbarium: yes no
STATUS: cultivar weed wild / market farm institute ( )/
landrace pure-line mutant improved( ) / lowland rainfed upland
LOCALITY: / km of
altitude: m
CROP SEASON:
CULTURAL PRACTICES:
USAGE:
NOTES
1 DISEASE and PESTS:
2 TOPOGRAPHY: swamp flood-plain plain undulating hilly mountainous
3 SITE: level slope summit depression 4 TEXTURE and STONINESS:
5 DRAINAGE: poor moderate good excess 6 ASSOCIATED PLANTS:
7 CHARACTERISTICS: plant height panicle length panicle num.
grains num. awn shattering color(culm hull
seed awn and apic.),
8 FARMER'S NAME and ADDRESS:
9 PHOTO NO. etc.:
10 ACCESSION NO. COL.NO. LOCAL NAME:

NIAR, MAFF COL.NO.	NIAR, MAFF COL.NO.	NIAR, MAFF COL.NO.	NIAR, MAFF COL.NO.

Fig. 2 An example of collection form adapted for gramineous cereal crops. This form is corresponding to IBPGR 'General form'.

LIST OF COLLECTED MATERIALS (_____) No._____

Collection No.	Date Month	Genus & Species	Cultivar or local name	Sample P/Ia[1]	Status[2]	Locality (Prov. Vill., km) & Altitude (m)	Crop season	Cultural practice	Usage	Diseases & pests	Topography[3]	Site[4]	Drainage[5]	Characteristics and others	Notes Name & address, etc.
1															
5															
10															
15															
20															

Note: Fill up in prints and arrange at every plant group. 1) Sample: P = population, In = individual sample, 2) Status: ① wild ② weedy ③ landrace ④ improved ⑤ breeder's line ⑥ others
3) Topography: ① swamp ② flood plain ③ plain level ④ undulating ⑤ hilly ⑥ mountainous ⑦ others. 4) Site: ① level ② slope ③ summit ④ depression.
5) Drainage: ① poor ② moderate ③ good ④ excessive

Fig. 3 A summarizing table of collected materials