UNIVERSIDADE DE SÃO PAULO
ESCOLA SUPERIOR DE AGRICULTURA "LUIZ DE QUEIROZ"
INSTITUTO DE GENÉTICA
PIRACICABA - S. Paulo - Brasil

Nº 286-USP/IGen
EXP/ES

June 21, 1966.

Dr. William L.Brown
Vice President, Director of Research
Pioneer Hi-Bred Corn Company
1206 Mulberry Street
Des Moines, Iowa 50308
U.S.A.

Dear Dr. Brown,

Thanks for your recent letter. By this time you should have already received a letter from Almiro about the International Genetics Symposium. The only thing I will say about that subject is that you may find the weather here a bit cold at night (the low usually is about 40º-50º), but the days are usually pleasant to cool. The early morning shave and shower often leaves something to be desired at first, but I have grown relatively accustomed to it now. The houses, hotels, and buildings are not heated and the only hot water usually is an electrically heated shower.

I guess I have been rather negligent about explaining what I'm trying to accomplish here. Briefly my ultimate goal is to develop some way to efficiently use the published morphological data to establish a tentative "family tree" of the races of maize. This is a rather far reaching goal, and in the end I may only demonstrate how not to accomplish this.

I am presently concerned more with the morphological data. (I have an idea about how to use the knob data, but as of yet have not tried to develop a method which uses both types of data. The method I have in mind for the knob data has been used in only one case so far. Some Italian workers used allelic frequencies in the blood groups of 15 human populations from almost all the four corners of the earth. The results look satisfactory in that case, and there seems to be no fundamental difference in the case of knob frequencies in maize.) There are various techniques available for the use of morphological data. I have high hopes that a very simple technique will yield logical results. The basic premise is that morphological similarity is evidence of close relationship. Thus I would like to have you consider the distribution of points in n-dimensional space, where each point represents a race of maize, and the coordinates of the points are the means of n morphological characteristics. The distances between these points (or race means) could be used as measures of relationship. This is a very old idea which

is not without merit, but there are two drawbacks. First, the distances are very sensitive to the choice of characters. The use of two or more characters which are strongly correlated among the race means (such as plant height and number of leaves per plant or ear length and central spike length) is one of the worst forms of character weighting. The degree of such weighting is generally unknown (and often unsuspected and undetected as well). Thus it is desirable to either use uncorrelated characters or to use a transformation to eliminate the effects of correlation. In general it is easier to transform. Since the distances are also very sensitive to choices of scale, the transformed variables must be standardized. In matrix form all this is very simple and the square of the distance between two transformed and standarized race means is

$$D_{12}^2 = (\underline{X}_1 - \underline{X}_2)' \; S^{-1} \; (\underline{X}_1 - \underline{X}_2)$$

where $\underline{X}_i$ is the vector of means for race i and S is the covariance matrix, estimated from the variation among as many as possible of the means of the races of maize. The distances between race boundaries are of course different than the distance between race means. The distance between race boundaries depends also upon the variation within races. The problem pertinent to the distances between race boundaries is whether the distributions of the plants (or, if you prefer, of the plot means) within races overlap. However, that is quite a different problem which seems to be of decidedly secondary importance at the moment.

The morphological data published for the races of maize seems sufficient to estimate the covariance matrix, although this data is far from ideal. This data will also supply the means, but there will be a serious problem with missing data. However, rather than beginning on such a grand scale, it seemed wise to first test the system with some populations of fairly well-known history, which represent a wide range of possible relationships. Since another system of measuring degree of relationship has been used here (with orchids) and is widely known, but infrequently used (with good reason in my opinion), I decided to collect data sufficient to contrast the results of the two techniques. For this purpose, individual plant data or at least replicated plot mean data were needed for each race. Although satisfactory data might have been located either here or elsewhere, it seemed more advisable (considering both the problems always present with stale data and my inexperience with the collection of races of maize) to collect new data.

Last summer I grew 15 races or sub-races from the collection here in a randomized complete block experiment with 8 replications. Individual plant data were taken on, wherever possible, 9 competitive plants per plot for plant height; leaf number, lenght, and width; number of leaves above the ear; number of primary tassel branches; peduncle,

UNIVERSIDADE DE SÃO PAULO
ESCOLA SUPERIOR DE AGRICULTURA "LUIZ DE QUEIROZ"
INSTITUTO DE GENÉTICA
PIRACICABA - S. Paulo - Brasil

branched part of tassel, and central spike lengths; ear lenght and diameter; kernel lenght, width, and thickness; row number; and maturity. The plants are lost now, but I still have all the ears. In fact, I am still measuring them. The races used were Amarillo de Ocho, Avatí Morotí (two populations), Avatí Djakaira, Avatí Pichinga, Avatí Pichinga Ihú, Lenha, Caingang, Cateto (three populations), Dent Paulista, Cravo (Rio Grande Dent), and Crystal (two populations).

I am hoping to have a few results by the time of the Symposium, but the outlook is not bright. Hopefully however, I will be able to end all this arithmetic by spring (October-November). At that time, I would like to make an abrupt switch and begin to study the colonial and precolonial history of Brasil. Because of my backlog of work here and because of my not very reliable Portuguese, I have not as yet taken advantage of the library facilities and personal contacts available. The only thing I have discovered is that no one seems to know what type of maize was being grown where when the Portuguese first arrived. Certain types are known to have been associated with certain Indian tribes, but there is no general agreement on exactly where the various Indian tribes were.

Don't be too alarmed by all the statistics, I am convinced that they can be helpful to interpret large volumes of data. However, any "family tree" based solely on statistical analyses of data will need to be carefully studied and revised in the light of other information. Nevertheless, almost any reasonable system of classification would seem to be better than the present situation. My impression is that any large scale usage of chromosome knob data is at least five years away. I have not yet studied any of the knob data carefully. I was impressed with the quantity of Kato's data, but personally would have prefered fewer samples of relatively purer races distributed geographically a bit more randomly.

Recently one of the assistants here (Dr. Maria Ruth Alleoni) asked me to help her obtain seed of the alleles at the Tu locus. She is trying to obtain the full set of alleles to use for testing stocks which she already has. In any event, she especially wants a stock with the tu^h allele. It appears that due to a long standing personal disagreement, there is not much chance that Mangelsdorf would send the necessary stocks. I find this a bit difficult to believe, but I don't know Mangelsdorf. However, as Gene Dalton has undoubtedly told you, Mangelsdorf does not hold me in very high regard. I would rather not open up any correspondence with him for fear of making him still more angry. However, I am deeply indebted to this particular assistant (she taught me about 80% of the Portuguese that I know and has been by far the most helpful person here). Thus if you could supply any of the necessary stocks she needs or suggest where it might be possible to obtain them, I would greatly appreciate it. (She has already sent a request to Galinat, but so far, at least, without results. It may be that Galinat is not in Massachusetts as I have been unable to obtain a reprint

UNIVERSIDADE DE SÃO PAULO
ESCOLA SUPERIOR DE AGRICULTURA "LUIZ DE QUEIROZ"
INSTITUTO DE GENÉTICA
PIRACICABA - S. Paulo - Brasil

of his recent article in Economic Botany despite several attempts since January).

Since Almiro and I are planning to go on another collecting trip in August, we would like to have you bring some water purification supplies with you if possible. There are at least three brands available: Halazone from Abbott Laboratories; Globaline from WTS Pharmaceuticals, Wallace and Trevrian, Inc. P.O.Box 1212, Rochester, N.Y. 14623 (1.98 for 50 tablets); and Potable-Aqua from Frost Laboratories, Inc. 430 Lexington Street, Auburndale, Mass. 02166 (1.65 for 50 tablets). This trip will be to the northern part of Mato Grosso, a region that probably even Coca Cola has not reached.

The Airport to use if possible is Viracopos. In any event it would be best to send us your flight number etc. well in advance. Best wishes for a good trip.

Sincerely,

Major

Major M.Goodman