Path Analysis of the Correlation between Fruit Number and Plant Traits of Cucumber Populations

Christopher S. Cramer[1]

Department of Agronomy and Horticulture, Box 30003, New Mexico State University, Las Cruces, NM 88003-8003

Todd C. Wehner[2]

Department of Horticultural Science, Box 7609, North Carolina State University, Raleigh, NC 27695-7609

Abstract. The relationships between fruit yield and yield components in several cucumber (*Cucumis sativus* L.) populations were investigated as well as how those relationships changed with selection for improved fruit yield. In addition, the correlations between fruit yield and yield components were partitioned into partial regression coefficients (path coefficients and indirect effects). Eight genetically distinct pickling and slicing cucumber populations, differing in fruit yield and quality, were previously subjected to modified half-sib family recurrent selection. Eight families from three selection cycles (early, intermediate, late) of each population were evaluated for yield components and fruit number per plant in four replications in each of two testing methods, seasons, and years. Since no statistical test for comparing the magnitudes of two correlations was available, a correlation (r) of 0.7 to 1.0 or –0.7 to –1.0 ($r^2 \geq 0.49$) was considered strong, while a correlation of –0.69 to 0.69 was considered weak. The number of branches per plant had a direct positive effect on, and was correlated ($r = 0.7$) with the number of total fruit per plant over all populations, cycles, seasons, years, plant densities, and replications. The number of nodes per branch, the percentage of pistillate nodes, and the percentage of fruit set were less correlated ($r < |0.7|$) with total fruit number per plant (fruit yield) than the number of branches per plant. Weak correlations between yield components and fruit yield often resulted from weak correlations among yield components. The correlations among fruit number traits were generally strong and positive ($r \geq 0.7$). Recurrent selection for improved fruit number per plant maintained weak path coefficients and correlations between yield components and total fruit number per plant. Selection also maintained weak correlations among yield components. However, the correlations and path coefficients of branch number per plant on the total fruit number became more positive ($r = 0.67$, 0.75, and 0.82 for early, intermediate, and late cycles, respectively) with selection. Future breeding should focus on selecting for the number of branches per plant to improve total fruit number per plant.

Fruit yield in cucumber is quantitatively inherited, with low heritability, making it difficult to improve (Smith et al., 1978). One way to improve yield indirectly is to select a second trait that has high heritability and is correlated with yield. The correlation between a particu-

lar trait and yield results from the direct effect of that trait and the indirect effects of other traits. Cramer (1997) reported that the number of branches per cucumber plant was positively correlated with total fruit number per plant in a pickling cucumber population. The positive association between total fruit and branch number resulted from the direct effect of branch number per plant and the indirect effect of other traits, such as node number per branch, percentage of pistillate nodes, and percentage of fruit set.

One goal of indirect selection is to find yield components that have a strong direct effect on fruit number per plant with little negative influence from other yield components. Path analysis can be employed to partition the correlation between yield components and yield into direct and indirect effects. Path analysis measures the direct effect of one trait on another trait by generating standardized partial regression coefficients (path coefficients) that are independent of

original units of measurement (Li, 1975; Wright, 1921, 1934). A large path coefficient indicates that the change will result in a proportional (or inversely proportional) change in another correlated trait, whereas a weak coefficient indicates that the change will have little effect on the second trait. In addition to the direct effects, the indirect effects of yield components should be calculated (based on the correlations among yield components). The indirect effect for each yield component is calculated by multiplying the correlation between two components by the direct effect of the opposite yield component. Path analysis has been used in a number of horticultural crops to study the relationships between yield and yield components (Hancock et al., 1984; Karlsson et al., 1988; McGiffen et al., 1994; Ranalli et al., 1981; Shasha'a et al., 1973; Siefker and Hancock, 1986).

We initiated several studies (Cramer and Wehner, 1998a, 2000) in cucumber to examine the potential of using selection of yield components to improve total fruit number per plant (total yield) indirectly. We investigated the correlations between yield components and total yield by partitioning the correlations into path coefficients and indirect effects through the use of path analysis. In addition, we examined the correlations among yield components, and the correlations among total, marketable, and early fruit number per plant. A major objective was to examine the influence of recurrent selection and testing method on correlations, path coefficients, and indirect effects.

Materials and Methods

Eight genetically distinct cucumber (four pickling and four slicing) populations that differed in fruit yield and quality (Wehner, 1997, 1998a, 1998b) were evaluated. Those populations were developed by intercrossing, then selected for 6 to 10 cycles using modified half-sib recurrent selection to improve fruit yield, earliness, and shape (Wehner and Cramer, 1996a, 1996b). Half-sib families were selected in the spring season based on a simple weighted index (SWI), which was based on yield (70%) and quality (30%) traits (Wehner and Cramer, 1996a, 1996b). Three cycles of recurrent selection (earliest, intermediate, latest) were chosen from remnant seed of each population (Cramer, 1997; Cramer and Wehner, 1998a, 2000). For each of the 24 population–cycle combinations, eight families were chosen randomly and self-pollinated in the greenhouse.

S₁ families were tested using a randomized complete-block design arranged in a split split-plot with four replications in each of two seasons (spring, summer) in each of 2 years (1995, 1996), using two testing methods: hill (6450 plants/ha) or plot (64,500 plants/ha). Whole plots were the eight cucumber populations (wide, medium, elite, and special of pickling and slicing types), subplots were three cycles of recurrent selection (early, intermedi-

ate, late), and sub-subplots were testing method (hill or plot) (Cramer, 1997; Cramer and Wehner, 1998a, 2000).

The planting methods and cultural practices have been reported previously (Cramer, 1997; Cramer and Wehner, 1998a, 2000). The plot testing method was chosen based on the plant density suggested by Wehner (1986) for pickling cucumbers grown for multiple harvest in North Carolina. The 3.1-m-long, bordered, test plots were harvested once-over by hand at the 10% oversize (>51 mm in diameter for pickles and >60 mm in diameter for slicers) stage as recommended by Miller and Hughes (1969) for optimum yield. Each plot was evaluated for fruit shape; total number of branches (main stem and lateral branches); number of nodes on main and lateral branches; pistillate flowers; and number of total, early, and marketable fruit. The number of branches per plant, number of nodes per branch, percentage of pistillate nodes, and percentage of fruit set were classified as yield components, while total, marketable, and early fruit number per plant were classified as fruit yield traits (Cramer, 1997).

Plots having a low stand were corrected to 30 plants per plot by dividing the observed yield component and fruit yield measurements by the current number of plants per plot and then multiplying by 30 (Cramer, 1997; Cramer and Wehner, 1998a, 1998b, 2000). Path coefficients, correlations among yield components and among yield traits, indirect effects of other yield components, and correlations between yield components and fruit yield were determined using PathSAS (Cramer et al., 1999). The PathSAS program standardizes the data to conform to a normal distribution with a mean of zero and a standard deviation of one. After standardization, multiple regression analysis is used to calculate the partial regression coefficients (path coefficients) for each yield component and its effect on total fruit number per plant. Pearson correlation coefficients were calculated for the relationships among yield components and among yield traits. Indirect effects were calculated by multiplying the respective correlation coefficient between the two yield components by the path coefficient for the indirectly related yield component. The correlation between a yield component and total yield was determined by summing the path coefficient between total yield and the yield component of interest with the indirect effects of the other yield components on total yield.

Since no statistical test for determining the importance of path coefficients was available, we decided to identify them as strong [0.7 to 1.0 or −0.7 to −1.0 ($r^2 \geq 0.49$)] or weak (−0.69 to 0.69). Correlations were calculated among yield components, among yield traits, and between yield components and total yield, for each selection cycle, and over testing methods, cycles, seasons, years, replications, and populations. In addition, estimates were made of path coefficients for each yield component and of the indirect effects of each yield component.

Results and Discussion

Latest cycle of selection. When averaged over all populations, cycles, seasons, years, plant densities, and replications, the number of branches per plant had a positive direct effect ($r = 0.74$) upon the total fruit number per plant (Fig. 1). As a result, the correlation between branch number and the total number of fruit per plant was strong and positive ($r = 0.70$) (Table 1). The strong association between branch number per plant and total fruit number per plant can be attributed partially to the use of LJ 90430 (*C. sativus* var. *hardwickii* Alef. PI 183967). That accession was used as a parent in the formation of several populations (Wehner, 1997, 1998b). LJ 90430 is unusual in having many branches and fruit per plant (Fredrick and Staub, 1989; Kupper and Staub, 1989; Staub, 1985). A strong correlation between branch and fruit number also was observed in a cross between two diverse cucumber lines (Serquen et al., 1997). One of the

parents used in their cross, H-19, had a large number of small lateral branches per plant, but did not have *C. sativus* var. *hardwickii* germplasm. When these lines were tested in Georgia, Serquen and co-workers (1997) identified a single quantitative trait locus (QTL) that explained 40% of the variation in lateral branch number and 9% of the variation in fruit number. The large contribution provided by a single QTL to lateral branch number may partially account for our finding of a strong correlation between the number of branches and total fruit per plant.

Branch number also exhibited a positive direct effect that was positive on marketable and early fruit number per plant. However, the effect was weaker than for total fruit number (Fig. 1, Table 1). The reduced direct effect of branch number on early fruit number per plant may be partially attributable to LJ 90430. In order for LJ 90430 to produce many fruit, plants must accumulate a large photosynthetic area to support the growth of a large number of
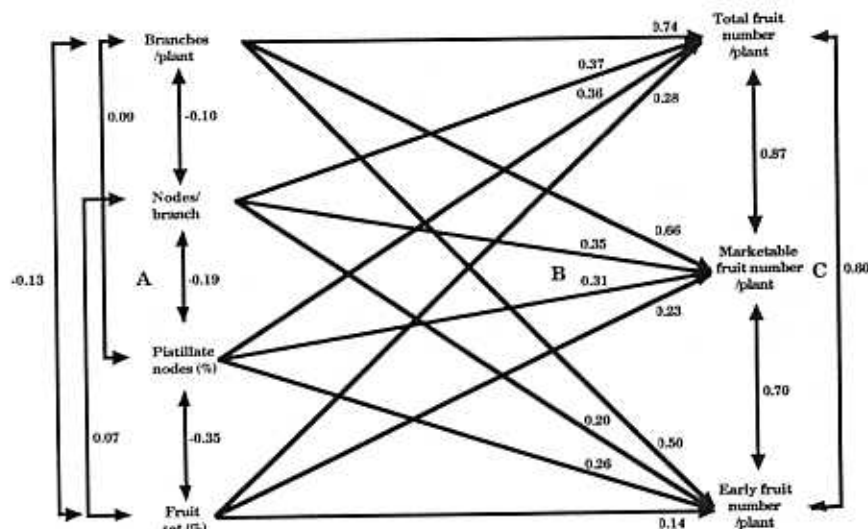

Fig. 1. (A) Path diagram of correlation coefficients among yield components, (B) direct effects of yield components on fruit number traits, and (C) correlation coefficients among fruit number traits in cucumber.

Table 1. Direct (underlined) and indirect effects and total correlations of yield components with total, marketable, and early fruit number per cucumber plant.

Yield component	Branches per plant	Nodes per branch	Pistillate nodes (%)	Fruit set (%)	Total correlation
Total fruit number per plant					
Branches per plant	0.74	−0.04	0.03	−0.04	0.70
Nodes per branch	−0.08	0.37	−0.07	0.02	0.25
Pistillate nodes (%)	0.07	−0.07	0.36	−0.10	0.25
Fruit set (%)	−0.10	0.03	−0.13	0.28	0.09
Marketable fruit number per plant					
Branches per plant	0.66	−0.04	0.03	−0.03	0.63
Nodes per branch	−0.07	0.35	−0.06	0.02	0.24
Pistillate nodes (%)	0.06	−0.07	0.31	−0.08	0.22
Fruit set (%)	−0.09	0.02	−0.11	0.23	0.06
Early fruit number per plant					
Branches per plant	0.50	−0.02	0.02	−0.02	0.49
Nodes per branch	−0.05	0.20	−0.05	0.01	0.11
Pistillate nodes (%)	0.05	−0.04	0.26	−0.05	0.21
Fruit set (%)	−0.06	0.01	−0.09	0.14	0.00

709

fruit. The small photosynthetic area of *C. sativus* var. *sativus* allows the plant to support only one developing, seeded fruit at a time (Pharr et al., 1985). LJ 90430 plants may circumvent the problem of "first-fruit inhibition" by having separate vegetative and reproductive growth phases. In doing so, the plants do not produce fruit early (<45 d) as do *C. sativus* var. *sativus* plants. Hence, the effect of branch number on early fruit number of *C. sativus* var. *hardwickii* is weaker than its effect on total fruit number of LJ 90430 plants. This trait has been maintained in some *C. sativus* var. *hardwickii*-derived cucumber populations.

Branch number was highly and positively ($r = 0.7$) correlated with total fruit number per plant suggesting that direct selection for branch number might result in larger gains in fruit yield than direct selection for fruit yield itself, particularly if branch number has higher heritability than fruit yield. Although heritability of branch number was not determined in our study, at least four QTLs were associated with variation for the number of lateral branches in a narrow cross using two diverse cucumber lines depending on the growing environment (Serquen et al., 1997). Therefore, the number of branches per plant would be expected to have higher heritability than fruit number, which is controlled by many genes and has low heritability (Smith et al., 1978).

When averaged over all populations, cycles, seasons, years, plant densities, and replications, all yield components besides branch number per plant exhibited weak ($-0.69 \le r \le 0.69$) direct effects and weak ($-0.69 \le r \le 0.69$) total correlations with total, marketable, and early fruit number per plant (Table 1). Solanki and Shah (1989) also observed weak correlations between the number of fruit per plant and the number of leaves (i.e., nodes per branch) and female flowers per plant (i.e., percentage of pistillate nodes). These results suggest that, without prior testing of the correlations between fruit number and yield component within a cucumber population, indirect selection based on the number of nodes per branch, the percentage of pistillate nodes, or the percentage of fruit set would not be recommended for increasing the number of cucumber fruit per plant. However, for particular population–cycle–season combinations, correlations between those yield components and total fruit number per plant can be strong and positive, and could be used in a breeding program to improve fruit number based on indirect selection (Cramer, 1997; Cramer and Wehner, 1998a, 2000). For example, the percentage of fruit set was positively correlated ($r = 0.80$) with the number of total and marketable fruit per plant of a wide-based slicer population (NCWBS) tested in the spring season in North Carolina (Cramer and Wehner, 1998a).

The weak correlations ($-0.69 \le r \le 0.69$) between the number of nodes per branch, the percentage of pistillate nodes, and the percentage of fruit set vs. total, marketable, and early fruit number per plant resulted from weak path coefficients ($-0.69 \le r \le 0.69$) for each yield component, and also resulted from weak

($-0.69 \le r \le 0.69$) indirect effects with other yield components (Table 1). The weak indirect effects probably resulted from the weak correlations among yield components (Fig. 1). The indirect effects on branch number per plant exhibited the lowest average absolute correlation among the four yield components (Table 1). The strong, positive correlation ($r = 0.7$) between branch and total fruit number per plant resulted from the strong, direct effect ($\rho = 0.74$) of branch number and the correlation was little influenced by other yield components. Hence, indirect selection based on branch number would not be influenced by other yield components, such as the number of nodes per branch. Although the number of branches per plant exhibited a strong path coefficient ($r = 0.74$) with total fruit number per plant, the indirect effects of branch number on total fruit number through other yield components was weak ($-0.69 \le r \le 0.69$), mainly as a result of the weak correlations ($r = -0.04, 0.03, -0.04$, respectively) between branch number and other yield components (Table 1). As a result, the indirect effects of branch number do not contribute to the correlation of other yield components with total fruit number per plant.

In our study, the correlations among fruit number traits were strong and positive ($r > 0.69$) when averaged over all populations, cycles, seasons, years, plant densities, and replications. Total and marketable fruit number per plant exhibited the strongest correlation ($r = 0.87$) among fruit number traits (Fig. 1). In previous studies strong, positive correlations between total and marketable fruit number per plant have been observed in pickle and slicer populations (Cramer, 1997; Cramer and Wehner, 1998a, 2000). In general, selection to increase the total number of fruit per plant also would increase the number of marketable and early fruit per plant.

Cycles of selection. In order to determine the effect of recurrent selection on yield components and fruit yield, direct effects, indirect effects, and correlations of these variables were calculated for each selection cycle (early, intermediate, late) averaged over populations, seasons, years, plant densities, and replications (Fig. 2, Table 2). Three of four of the path coefficients and correlations between yield components and total fruit number per plant were weak at the early selection cycle and remained weak with selection. Of the four yield components, only the number of branches per plant had a strong, positive path coefficient ($\rho = 0.75$) and correlation ($r = 0.72$) with the total number of fruit per plant at the intermediate cycle of selection, and this strong correlation was increased with selection ($r = 0.77$) (Fig. 2, Table 2). The contribution of *C. sativus* var. *hardwickii* to those populations may account for the strong association between branch and fruit number. Weak correlations among yield components at each cycle of selection (Fig. 2) resulted in weak indirect effects for each yield component at each cycle of selection (Table 2). Even though the direct effects of branch number per plant increased with selection, the indirect effects of branch number on other yield components were weak initially and remained weak with selection. The weak indirect effects resulted from weak correlations between branch number and the other yield components. Thus, the number of branches per plant could be selected to improve fruit yield without changing the relationship between branches per plant and other yield components.

The correlations among total, marketable, and early fruit number per plant also were examined to determine the effects of selection on the relationships among these traits. The correlation between total and marketable fruit number per plant was positive and strong
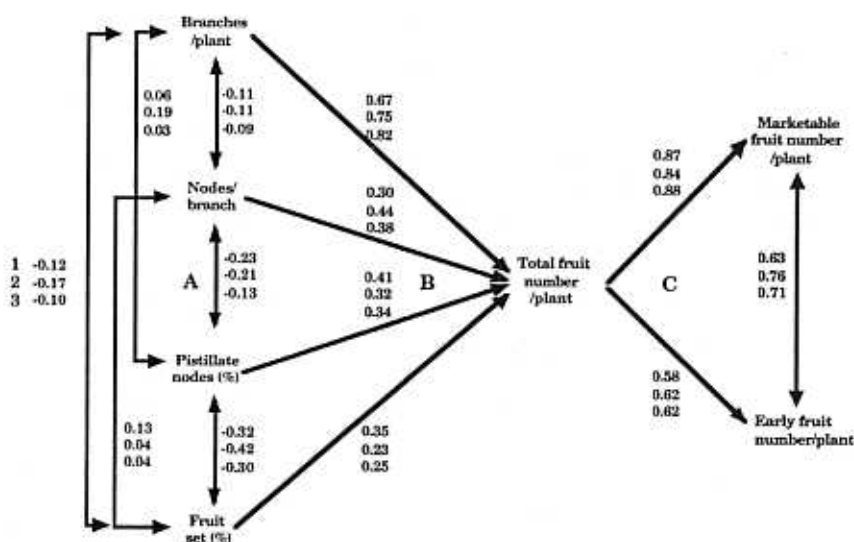

Fig. 2. (A) Path diagram of correlation coefficients among yield components, (B) direct effects of yield components on total fruit number per plant, and (C) correlation coefficients among fruit number traits in cucumber for three cycles of selection [early (1), intermediate (2), late (3)].

Table 2. Direct (underlined) and indirect effects and total correlations of yield components with total fruit number per cucumber plant during recurrent selection (three cycles).

Yield component	Branches per plant	Nodes per branch	Pistillate nodes (%)	Fruit set (%)	Total correlation
		Branches per plant			
Early	0.67	−0.03	0.02	−0.04	0.62
Intermediate	0.75	−0.05	0.06	−0.04	0.72
Late	0.82	−0.04	0.01	−0.03	0.77
		Nodes per branch			
Early	−0.08	0.30	−0.10	0.05	0.17
Intermediate	−0.08	0.44	−0.07	0.01	0.30
Late	−0.08	0.38	−0.05	0.01	0.27
		Pistillate nodes (%)			
Early	0.04	−0.07	0.41	−0.11	0.26
Intermediate	0.14	−0.09	0.32	−0.10	0.27
Late	0.02	−0.05	0.34	−0.07	0.24
		Fruit set (%)			
Early	−0.08	0.04	−0.13	0.35	0.18
Intermediate	−0.13	0.02	−0.13	0.23	−0.01
Late	−0.08	0.01	−0.10	0.25	0.08

($r = 0.87$) at the early cycle and remained strong with selection ($r = 0.88$ at the latest cycle) (Fig. 2). The correlation between total and early fruit number per plant was positive and weak ($r = 0.58$) at the early cycle and remained weak ($r = 0.62$) (e.g., at the latest cycle). The correlation between early and marketable fruit number per plant was weak and positive at the initial cycle of selection ($r = 0.63$) and became strongly positive by the latest cycle of selection ($r = 0.71$). Thus, selection to increase the number of total fruit per plant also should increase the number of marketable fruit per plant. In addition, selection to increase the number of marketable fruit per plant will increase the number of total and early fruit per plant.

For the yield components measured, the number of branches per plant had the greatest direct effect on total fruit number per plant; correlations between other yield components and total fruit number were weak. Moreover, the direct effects of the number of nodes per branch, percentage of pistillate nodes, and percentage of fruit set on the total number of fruit per plant were weak. Likewise, with respect to total fruit number per plant, the indirect effects of yield components on each other were weak for all yield components. Those weak indirect effects and correlations resulted mainly from weak correlations among yield components. Therefore, selection for an increased number of branches per plant could be used to improve total yield, whereas selection for number of nodes per branch, percentage of pistillate nodes, or percentage of fruit set would not increase fruit yield. Selection for

an increase in the number of branches per plant will result in little change in other yield components.

Literature Cited

Cramer, C.S. 1997. Specific combining ability for fruit yield and shape, yield and yield components of cucumber (Cucumis sativus L.) populations improved using recurrent selection. PhD Diss., North Carolina State Univ., Raleigh.

Cramer, C.S. and T.C. Wehner. 1998a. Fruit yield and yield component means and correlations of four slicing cucumber populations improved through six to ten cycles of recurrent selection. J. Amer. Soc. Hort. Sci. 123:388–395.

Cramer, C.S. and T.C. Wehner. 1998b. Stand correction methods in cucumber. Cucurbit Genet. Coop. Rpt. 21:18–20.

Cramer, C.S. and T.C. Wehner. 2000. Fruit yield and yield component correlations of four pickling cucumber populations. Cucurbit Genet. Coop. Rpt. 23: (In press.).

Cramer, C.S., T.C. Wehner, and S.B. Donaghy. 1999. PathSAS: A SAS computer program for path coefficient analysis of quantitative data. J. Hered. 90:260–262.

Fredrick, L.R. and J.E. Staub. 1989. Combining ability analyses of fruit yield and quality in near-homozygous lines derived from cucumber. J. Amer. Soc. Hort. Sci. 114:332–338.

Hancock, J.F., M.P. Pritts, and J.H. Siefker. 1984. Yield components of strawberries maintained in ribbons and matted rows. Crop Res. 24:37–43.

Karlsson, M.G., M.P. Pritts, and R.D. Heins. 1988. Path analysis of chrysanthemum growth and development. HortScience 23:372–375.

Kupper, R.S. and J.E. Staub. 1989. Combining ability between lines of Cucumis sativus and Cucumis sativus var. hardwickii Alef. Euphytica 38:197–210.

Li, C.C. 1975. Path analysis—A primer. Boxwood Press, Pacific Grove, Calif.

McGiffen, M.E., Jr., D.J. Pantone, and J.B. Masiunas. 1994. Path analysis of tomato yield components in relation to competition with black and eastern black nightshade. J. Amer. Soc. Hort. Sci. 119:6–11.

Miller, C.H. and G.R. Hughes. 1969. Harvest indices for pickling cucumbers for once-over harvested systems. J. Amer. Soc. Hort. Sci. 94:485–487.

Pharr, D.M., S.C. Huber, and H.N. Sox. 1985. Leaf carbohydrate status and enzymes of translocate synthesis in fruiting and vegetative plants of Cucumis sativus L. Plant Physiol. 77:104–108.

Ranalli, P., M. DiCandilo, I. Giordano, and B. Casarini. 1981. Correlation and path analysis in peas (Pisum sativum L.) for processing. Z. Pflanzenzuecht. 86:81–86.

Serquen, F.C., J. Bacher, and J.E. Staub. 1997. Mapping and QTL analysis of horticultural traits in a narrow cross in cucumber (Cucumis sativus L.) using random-amplified polymorphic DNA markers. Mol. Breeding 3:257–268.

Shasha'a, N.S., W.P. Nye, and W.F. Campbell. 1973. Path coefficient analysis of correlation between honey bee activity and seed yield in Allium cepa L. J. Amer. Soc. Hort. Sci. 98:341–347.

Siefker, J.H. and J.F. Hancock. 1986. Yield component interactions in cultivars of the highbush blueberry. J. Amer. Soc. Hort. Sci. 111:606–608.

Smith, O.S., R.L. Lower, and R.H. Moll. 1978. Estimates of heritabilities and variance components in pickling cucumbers. J. Amer. Soc. Hort. Sci. 103:222–225.

Solanki, S.S. and A. Shah. 1989. Path analysis of fruit yield components in cucumber. Progressive Hort. 21:322–324.

Staub, J.E. 1985. Preliminary yield evaluation of inbred lines derived from Cucumis sativus L. var. hardwickii (Royle) Kitamura. Cucurbit Genet. Coop. Rpt. 8:18–21.

Wehner, T.C. 1986. Efficiency of 3 single-harvest tests for evaluation of yield in pickling cucumber. Euphytica 35:493–501.

Wehner, T.C. 1997. Three pickling cucumber populations: NCWBP, NCMBP and NCEP1. HortScience 32:941–944.

Wehner, T.C. 1998a. Three slicing cucumber populations: NCWBS, NCMBS and NCES1. HortScience 33:168–170.

Wehner, T.C. 1998b. Two special cucumber populations: NCH1 and NCBA1. HortScience 33:766–768.

Wehner, T.C. and C.S. Cramer. 1996a. Gain for pickling cucumber yield and fruit shape using recurrent selection. Crop Sci. 36:1538–1544.

Wehner, T.C. and C.S. Cramer. 1996b. Ten cycles of recurrent selection for fruit yield, earliness, and quality in three slicing cucumber populations. J. Amer. Soc. Hort. Sci. 134:322–326.

Wright S. 1921. Correlation and causation. J. Agr. Res. 20:557–585.

Wright S. 1934. The method of path coefficients. Ann. Math. Stat. 5:161–215.