Low Correlation Between Genomic and Morphological Introgression Estimates in a Walnut Backcross

Keith Woeste¹ and Gale McGranahan

Department of Pomology, Wickson Hall, University of California, Davis, CA 95616

Robert Bernatzky

Department of Plant and Soil Science, French Hall, University of Massachusetts, Amherst, MA 01003

ADDITIONAL INDEX WORDS. *Juglans regia*, *Juglans hindsii*, Paradox, Mahalanobis' distance

ABSTRACT. A first backcross population of walnuts [*Juglans hindsii* (Jeps.) Jeps. × *Juglans regia* L.] × *J. regia*] was used to evaluate the correlation between morphological (statistical) and genetic distance during introgression. Five traits based on leaf morphology were identified to quantitate the morphology of the parental species, their F₁ hybrids, and the backcrosses to each parent. These traits were used to evaluate the morphological similarity of first backcrosses to *J. regia* using Mahalanobis' distance. The amount of genomic introgression of each backcross was estimated using 59 randomly amplified polymorphic DNA (RAPD) and 41 restriction fragment-length polymorphism (RFLP) genetic markers that identify polymorphisms between *J. regia* and *J. hindsii*. A smaller scaffold set of markers was also identified using published linkage data. The correlation between the measures of morphological and genomic introgression for the first backcrosses was low (0.23) but significant. The results suggest that selection based on morphology during backcrossing will not be an effective way to recover recurrent parent genome.

The morphological evaluation of populations has been used by geneticists and taxonomists to characterize introgression. The idea that morphological intermediacy in wild populations implies genetic introgression was introduced by Anderson (1949) and Fisher (1936). Since their path-breaking work, a variety of statistical distance measures have been used to study introgression, evaluate hybrid complexes, infer parentage or group classification, and estimate genetic distance (Brochman, 1987; Ietswaart and Feij, 1989; Namkoong, 1966; Yang and Selander, 1968; Zeyl and Lowcock, 1989). Mahalanobis' distance (MD) (1936), a multivariate method for evaluating distance, is particularly appropriate for analyzing introgression because it removes the correlations among traits. It is the preferred method of analysis when the variance and covariance structures of the populations to be evaluated are unequal (Atchley et al., 1982a; Beer et al., 1993). Plant breeders have used the same distance measures to identify and classify germplasm and evaluate its genetic variability (Baum and Bailey, 1984; Goodman, 1968; Pereira-Lorenzo et al., 1996; Souza and Sorrels, 1991; Smith and Smith, 1989a). Plant breeders rarely use statistical distance measures to evaluate introgression, since replicating observations of many traits is expensive and the goal of most introgression programs is to backcross as quickly as possible.

Ecologists and plant breeders have also used genetic markers such as isozymes, randomly amplified polymorphic DNA (RAPD) markers, and restriction fragment-length polymorphisms (RFLPs) to evaluate genetic distance and predict the potential effects of combining gene pools (Dudley, 1994; Cowen and Frey, 1987; Kesseli et al., 1991; Moser and Lee, 1994; Smith and Smith, 1989b; Williams and St. Clair, 1993). Openshaw et al. (1994),

Tanksley et al. (1981), and others have proposed that plant breeders could use molecular markers to reduce linkage drag and increase the efficiency of introgression. One method, outlined by Hillel et al. (1990) and Hospital et al. (1992), uses genetic markers to monitor genome introgression by selecting against the entire donor genome except a particular trait of interest. This approach has been called genomic selection (Hillel et al., 1990).

The Persian walnut (*Juglans regia* L.), and the Northern California black walnut [*Juglans hindsii* (Jeps.) Jeps.] are morphologically and genetically distinct, being members of different subsections of the genus *Juglans*. In this paper we apply a statistical distance measure (MD) and genomic selection to a first backcross to evaluate the introgression of Northern California black walnut into the cultivated Persian walnut. Because walnuts are long-lived trees with a long generation time, selection during backcrossing is potentially useful. In this paper we present data that indicates that statistical and genetic distance measures provide only weakly concordant estimates of the amount of introgression among first backcross trees.

Materials and Methods

CHOOSING TRAITS FOR MORPHOLOGICAL EVALUATION. We used the following criteria for choosing morphological traits for analysis of introgression: 1) they are easy to measure, 2) they have apparent quantitative inheritance, 3) the parental species must differ for the trait without overlap in range, 4) hybrids should be intermediate between their parents (Namkoong, 1966), 5) the traits should be vegetative and present in juvenile plants grown from seed to help evaluate heritability and early selection, and 6) the traits must not have been the target of selection. Identifying traits that met these criteria was difficult; after the preliminary evaluation of >10 traits we identified 5 that could be reliably and easily scored: the ratio of the subterminal leaflet length/subterminal leaflet width (SL/SW), the ratio of the rachis length/number of leaflets (RL/N), the ratio of the basal leaflet length/basal leaflet width (BL/BW), the ratio of the subterminal leaflet width/rachis length (SW/RL), and the ratio of the petiole length/rachis length (PL/RL).

Received for publication 11 Apr. 1997. Accepted for publication 22 Sept. 1997. We acknowledge the California Walnut Marketing Board for their support of this research. The cost of publishing this paper was defrayed in part by the payment of page charges. Under postal regulations, this paper therefore must be hereby marked advertisement solely to indicate this fact.

¹Former graduate research assistant. Currently postdoctoral researcher, Laboratory for Molecular Biology (M/C 567), Molecular Biology Research Facility, University of Illinois at Chicago, 900 South Ashland Ave., Chicago IL 60607. To whom reprint requests should be addressed.

Table 1. Walnut and walnut hybrid populations used to evaluate interpopulation morphological variability.

Population	Abbreviation	Location (Calif.)	Description
<i>Juglans regia</i>			
Persian walnut	JR	Davis, Winters	29 genotypes, including old and recent cultivars, breeding lines and germplasm
<i>Juglans hindsii</i>	JH	Davis, Winters	12 germplasm accessions and 26 wild genotypes
F ₁ interspecific hybrids (JH × JR)	F ₁	Davis	15 genotypes of unknown parentage collected as naturally occurring hybrids and selected for fruitfulness or as rootstocks
First backcross [(JH × JR) × JR]	BC ₁	Davis	Open-pollinated progeny of a single F ₁ selected for fecundity in crosses with <i>J. regia</i>
First backcross [(JH × JR) × JH]	BCH	Davis	15 genotypes from seed of an open-pollinated F ₁ selected for fecundity in crosses with <i>J. hindsii</i>

PLANT POPULATIONS AND MORPHOLOGICAL DATA COLLECTION. Mature *J. regia*, *J. hindsii*, their F₁ hybrids {Paradox [any F₁ hybrid between *J. regia* and *J. hindsii* is designated Paradox (McGranahan, personal communication)]}, BC₁s (F₁ × *J. regia*), BCH (backcrosses to *J. hindsii*, F₁ × *J. hindsii*), and juvenile second backcrosses to *J. regia* (BC₂) were characterized using the traits described above (Table 1). BC₁ was comprised of 84 random seedlings of a single, open-pollinated F₁ hybrid individual. All such F₁ hybrids are male sterile (Woeste et al., 1998). The plants described in this study were maintained by the Dept. of Pomology at the Univ. of California, Davis, according to current recommendations (Ramos, 1985; University of California Statewide Integrated Pest Management Project, 1982).

MEASUREMENT METHOD. On mature trees, morphological traits were measured to the nearest millimeter on randomly chosen, healthy appearing, fully expanded leaves that were exposed to full sun for at least part of the day. Leaves on shoots with actively growing tips were not measured. On juvenile plants, observations were made on healthy appearing, fully expanded leaves; usually the oldest leaves on the shoot for which precise measurements could be made. In all cases each observation is the mean of three measurements taken from three leaves per individual at the same time. Measured leaflets were always on the same side of the (pinnate compound) leaf. An effort was made to measure leaves from different sides of the canopy. Mature genotypes were evaluated for 2 to 3 years.

ESTIMATES OF GENETIC PARAMETERS. Estimates of the repeatability of the morphological traits were based on mature BC₁ trees evaluated in 1990–93. Neither all individuals nor all variables were evaluated in every year. To determine the significance of genotypic effects for the morphological traits, 18 randomly selected genotypes from each population were measured from 1990 through 1993. Repeatabilities and standard errors were calculated using the method of Becker (1985).

Estimates of the narrow-sense heritability (h^2) of the morphological traits were obtained using a regression of offspring on biparental means (Falconer, 1981). BC₂ progeny were in 54 full-sib families composed of an average of 3.3 seedlings per family. BC₂ was produced from crosses of *J. regia* 'Chandler' (used as a male) to the members of the BC₁, which was a mixture of half-sibs and full-sibs all derived from a single open-pollinated F₁. Biparental means were based on single observations and progeny values were means of single observations of all the progeny produced by a cross. An adjustment was made for differences in parent and offspring grand means by using deviations from trait means as dependent and independent variables (Fernando and Gianola, 1988). Standard errors for heritability were estimated as in Falconer (1981). All observations for heritability estimates were log₁₀ transformed before analysis. The offspring and parents were

grown in separate but similarly situated blocks in the Univ. of California, Davis, pomology fields.

MAHALANOBIS' DISTANCE. The statistical distance between two populations *i* and *j* was calculated as Mahalanobis' generalized distance Eq. [1]:

$$MD_{ij} = [(x_i - x_j)' S^{-1} (x_i - x_j)]^{1/2} / P \quad [1]$$

where *P* was the number of variables (5), x_i was the vector of character means for the reference population *i* (*J. regia*), and S^{-1} was the estimate of the common within population phenotypic covariance matrix for mature *J. regia* germplasm. The MD of any individual from the reference population was calculated as an individual's mean value for each of the traits during all the years data was collected. For some genotypes this mean was composed of many observations over several years. When between-population comparisons were made, MDs were log₁₀ transformed to make them conform more closely to the assumptions of ANOVA (Fernandez, 1992).

GENETIC MARKERS AND GENOMIC INTROGRESSION. Genomic introgression was evaluated in 76 BC₁s using 59 RAPD and 41 RFLP genetic markers that identify *J. regia* and/or *J. hindsii* loci polymorphic in the BC₁ (Woeste et al., 1996; Fjellstrom and Parfitt, 1994). The RFLP loci were identified by Fjellstrom (1993), and segregated as codominant markers in the BC₁. The RAPD markers were fixed (present) in *J. hindsii* but absent from *J. regia* (fixed null). The RAPDs segregated as present (in the heterozygous state) or absent in the BC₁. The markers are in 14 linkage groups covering about 1000 cM of the estimated 1600 cM genetic map of walnut. Only markers that did not deviate significantly from the expected 1:1 segregation in the BC₁ were used in the analysis (Woeste et al., 1996; Fjellstrom, 1993). The percent recipient parent genome (PRPG) of each BC₁ (Eq. [2]) was calculated as the proportion of the BC₁ genome derived entirely from *J. regia* (0.5) plus the product of the proportion of the genome segregating for *J. regia* and *J. hindsii* alleles (0.5) and the proportion of the loci homozygous for *J. regia* expressed as a percent.

$$PRPG = \{ [0.5 + (0.5 \times \text{no. homozygous loci/no. loci observed})] \times 100 \} \quad [2]$$

A data set of 100 markers and a subset of 50 RAPD and RFLP markers spaced ≈20 cM from the next marker were used to calculate PRPG. We refer to the set of 50 markers as a "scaffold" set. Selection on markers is theoretically more efficient when evenly spaced marker sets are used (Hospital et al., 1992). Not all markers were scored in every BC₁; an average of 67 markers per genotype was scored when the full data set was used, and 33.0 markers per genotype were scored when the scaffold data set was used.

STATISTICAL ANALYSES. The data were analyzed using SAS

edition 6.03 (Cary, N.C.). Simple descriptive statistics were calculated using the MEANS procedure. The RANK procedure was combined with the SPEARMAN option of the CORR procedure to generate rank correlations. ANOVA was run using GLM. Differences in the trait means for the parental species, for both juvenile and mature walnuts, were evaluated using Student's *t* test or REGWQ. To estimate the distribution of trait variance among and within populations, a sample of 7 mature trees of *J. regia* (cvs. Hartley, Howard, Payne, Pedro, Serr, Sunland, Tehama), BC₁ (vk9, 77, 80, 85, 86, 93, 95) and F₁ (BK, Burbank, Mom, PC, C101, Driver, Walt's) genotypes was used. The ANOVA for differences among BC₁ genotypes for the morphological traits was based on observations of 18 mature BC₁s (vk2, 13, 14, 22, 26, 30, 35, 36, 37, 38, 39, 40, 108, 110, 111, 114, 117 and 121) from 1990 to 1993.

Results

There were highly significant differences between *J. regia* and *J. hindsii* for the morphological traits for both mature and juvenile plants (in all cases $P \leq 0.0001$). The morphology of F₁ hybrids was intermediate for each trait, and the backcross means were intermediate between the F₁ and the parental species, although the population ranges overlapped extensively (data not shown). As the amount of *J. regia* genome in hybrids increased, leaflets were shorter and wider (more oblong and less lanceolate); there was greater space between leaflets, the ratio of leaflet width to leaf length increased and the distance from the base of the rachis to the first leaflet increased as a proportion of the total leaf length.

Variance for SL/SW, BL/BW, SW/RL and PL/RL was significant both among and within populations (Table 2). There was significant variance for RL/N only among populations. For SL/SW, SW/RL and PL/RL more than half of the variance was among populations, reflecting the ease with which mature members of these populations can be visually distinguished. The BC₁ population was significantly different from *J. regia* and the F₁ hybrids for all traits but BL/BW. The proportion of variance within populations due to differences among genotypes varied from 9% to 47%. Repeating measures over years as we did for this analysis does not

lead to strong conclusions since the variance for genotypes within populations was completely confounded with genotype $\times$ year interactions. Confounding of this type may have been significant for some of the populations but was not found to be so in an analysis that included only *J. regia* cultivars (data not shown).

We used estimates of genetic parameters to verify that the traits used to evaluate morphological introgression had a genetic basis. These parameters estimate the extent to which hybrid morphology for the traits represents genomic mixing and not simply random environmental effects. Repeatability (R) is the upper bound of the broad-sense heritability (Falconer, 1981). The repeatability of all the traits was moderately high in the BC₁, ranging from 0.53 for RL/N to 0.76 for BL/BW (Table 3). During the years of evaluation slightly over half of the differences among BC₁ genotypes for each of the traits were attributable to genotypic differences.

Narrow-sense heritability (h^2) is an estimate of the proportion of the total variance due to additive genetic effects. We estimated the narrow-sense heritabilities of the five morphological traits by regression of offspring values on biparental means (Table 3). The estimates were significant for SL/SW and SW/RL. Very high heritabilities were found for other characters in walnut based on plot means (Hansche et al., 1972). Heritabilities of 0.3 or 0.4 are commonly observed for morphological traits (Cheverud, 1988). Our estimates of the genetic parameters (Table 3) were certainly biased by the genetic structure of the populations used in the analysis, probably leading to overestimation of the parameters. For example, the estimate of the heritability of the trait SW/RL was larger than unity, the theoretical upper bound for heritability. This inaccuracy may be due to the small size of the populations used to generate the estimate, aberrant segregation in the backcrosses or selection during the breeding process. All of these factors were likely present in some BC₂ families since many of the female parents (BC₁s) had very low fertility and produced offspring with poor vigor. The heritability estimates were biased by several other factors as well: because we used a single male parent ('Chandler') and a BC₁ population derived from a single, open-pollinated female, the variance of the parents (the denominator in the heritability estimate) was reduced. Since the male parent(s) was presum-

Table 2. ANOVA for population effects for the morphological traits² used to calculate Mahalanobis' distance, the distribution of the variance among sources of variation, and mean separation of the populations.³

Trait	Source ⁴	df	Mean square	F	Variance (%)	Mean separation	
						JR:BC ₁	BC ₁ :F ₁
SL/SW	Population	2	0.1505	22.5***	56.7	*	***
	Genotype (population)	18	0.0067	5.15***	31.1		
	Error	63	0.0013		12.1		
RL/N	Population	2	0.0524	9.57**	31.5	*	**
	Genotype (population)	18	0.0055	1.77	14.8		
	Error	63	0.0031		33.7		
BL/BW	Population	2	0.0255	4.55*	26.1	NS	**
	Genotype (population)	18	0.0056	4.67**	47.2		
	Error	25	0.0012		26.7		
SW/RL	Population	2	0.9070	47.9***	77.6	***	***
	Genotype (population)	18	0.0189	4.12***	10.9		
	Error	63	0.0046		11.5		
PL/RL	Population	2	0.1143	42.3***	72.4	***	***
	Genotype (population)	18	0.0027	2.69*	9.0		
	Error	44	0.0010		18.6		

²See text for descriptions of the morphological traits.

³Populations were *Juglans regia* (JR), *J. regia* $\times$ *J. hindsii* (F₁), and the first backcross to *J. regia* (BC₁).

⁴Genotypes are nested within population.

NS, *, **, *** Nonsignificant or significant at $P = 0.05, 0.01, 0.001$, respectively. Population mean separation was by LSD.

Table 3. Heritability and repeatability estimates of the morphological traits used to calculate Mahalanobis' distance.

Trait ^a	$h^2 \pm SE^b$	$R \pm SE$
SL/SW	$0.45 \pm 0.19^*$	0.628 ± 0.044
RL/N	0.11 ± 0.19	0.530 ± 0.051
PL/RL	0.15 ± 0.19	0.589 ± 0.047
SW/RL	$1.25 \pm 0.19^{***}$	0.595 ± 0.054
BL/BW	0.41 ± 0.19	0.759 ± 0.116

^aSee text for description of the traits.

^bBased on 54 BC₂ families.

*.***Significant at $P = 0.05$ or 0.001 , respectively.

ably heterozygous for genes affecting morphology, the covariance of the offspring and parents (the numerator) was probably decreased. The relative importance of these biases in the estimates is not known.

MAHALANOBIS' DISTANCE AS A MEASURE OF INTROGRESSION. We used MD to create a single composite morphological score based on the five traits described above. MD was calculated as the morphological similarity of walnut genotypes or populations to the mean *J. regia* morphology and is thus the statistical measure of morphological introgression. The BC₁, F₁ and BCH hybrid populations represent a range of mixing of the *J. hindsii* and *J. regia* genomes; assuming normal segregation they contain 75%, 50%, and 25% *J. regia* genome, respectively. The more introgressed a population, the smaller its mean MD from *J. regia* (Table 4). As measured by MD, *J. regia*, the F₁ hybrids and *J. hindsii* are morphologically distinct, without overlap in range. *J. regia* was the most variable for MD and *J. hindsii* the least variable, but all the populations were significantly distinct morphologically. The distribution of the MD of the BC₁ is roughly bell-shaped with short tails and a skew toward the male (*J. regia*) parent (Fig. 1). The median and mean MD of the BC₁ were 1.35 and 1.34, respectively, after log transformation, with a standard deviation of 0.27. The most extreme MD values in the BC₁ overlapped with the range of MD of the parental types (*J. regia* and the F₁ hybrids). Thus, the majority of the BC₁s were morphologically intermediate between *J. regia* and the F₁s, but some of the BC₁s were morphologically indistinguishable from the parental types. Some of the variability of the BC₁ was environmental, since the mean MD of the BC₁ was significantly different in 1992 and 1993 based on Student's *t* test ($P = 0.013$). Despite this variability, there were significant differences for MD among BC₁ genotypes ($P \leq 0.001$) and the rank correlation between the MD of 30 mature BC₁ genotypes in 1992 and 1993 was 0.718 ($P \leq 0.0001$). The MD had a repeatability of ($R = 0.42 \pm 0.21$) based on observations of 30 BC₁ genotypes in 2 years.

The *J. regia* population was the most variable for MD as measured by cv, and *J. hindsii* the least variable (Table 4). At least

two factors probably contributed to the large difference in cv between the two parental populations: a diverse *J. regia* germplasm was used in the analysis whereas the germplasm base for *J. hindsii* is quite narrow, and the variables which comprise the MD may better describe the morphological variation within *J. regia* than *J. hindsii*. The variability for MD within the F₁ was quite low. The genetic variability within these hybrids is not well characterized and it is possible the F₁s used in the study are related since they were naturally occurring hybrids and only a handful of genotypes of *J. regia* are widely planted in California. Moreover it is not known if all or only a subset of *J. hindsii* and *J. regia* genotypes can hybridize.

PERCENT RECIPIENT PARENT GENOME. The percent recipient parent genome (PRPG) is an estimate of genomic introgression based on the segregation of genetic markers derived from *J. hindsii* and *J. regia* in the BC₁. The PRPG of the BC₁ ranges from a high of 87.1% recipient parent genome to a low of $\approx 65\%$ with a mean of 74% (Fig. 2) approximating the theoretically expected distribution (Hillel et al., 1990). Although the distribution appears roughly symmetric, there were more heterozygous than homozygous loci in the BC₁ (1297 versus 1189) and the difference was significant when tested against an expected 1:1 segregation ($\chi^2 = 4.6$, $P \leq 0.05$) (Woeste, 1996). This difference skewed the PRPG slightly toward the F₁ (female) parent. The standard deviation of the PRPG (5.0%) is also slightly higher than the 3.6% predicted by Hillel et al. (1990) for first backcrosses. The PRPG of the BC₁ was also calculated using a scaffold set of 50 loci. Rank correlations between the larger and smaller data sets were high ($r = 0.72$) and significant (Table 5).

COMPARISON OF MD AND PRPG. Although the MD and the PRPG each produce reasonable descriptions of the introgression in the BC₁, the correlation between the two methods while statistically significant is strikingly low ($r = 0.23$; Table 5). Backcrosses that were statistically quite similar to *J. regia*, i.e., that were virtually identical to *J. regia* morphologically, often contained as much as 25% *J. hindsii* genome. The rank correlation between MD and PRPG based on a scaffold set of markers was also quite low ($r = 0.14$) and nonsignificant. In our population two single morphological traits SL/SW and SW/RL, had equal or better correlations with PRPG than the MD (Table 5).

Discussion

In this study the PRPG and MD were investigated as measures of linkage drag. These measures could then be used for selecting among backcrosses. The results demonstrate that the parental species and their hybrids have significantly different morphologies. The estimates of genetic parameters, as well as the significant differences for the traits between and within populations, suggest that the observed morphological differences have a genetic basis. When combined into a single, statistical measure of introgression,

Table 4. Statistical summary and means separation of the log-transformed Mahalanobis' distance (MD) from mature *Juglans regia* to individual, mature *J. regia*, *J. hindsii*, F₁ hybrids of these species and their backcrosses.

Population	n ^a	MD				
		Minimum	Maximum	Mean ^b	SD	CV (%)
<i>Juglans regia</i>	24	-0.08	1.18	0.60 a	0.30	50
Backcross (BC ₁)	80	0.64	1.84	1.34 b	0.27	20
F ₁ hybrids	11	1.57	1.91	1.76 c	0.10	5.7
Backcross (BCH)	14	1.82	2.34	2.07 d	0.16	7.7
<i>Juglans hindsii</i>	15	2.11	2.49	2.32 e	0.09	3.9

^aSample size is the number of genotypes evaluated in each population, not the number of observations.

^bMeans followed by the same letter are not significantly different, $P \leq 0.05$, by REGWQ test.

Fig. 1. Distribution of Mahalanobis' distance (MD) for 80 first backcross walnuts. Arrows indicate mean MD for *Juglans regia* and *J. hindsii* × *J. regia* F₁ populations.

Fig. 2. Distribution of percent *Juglans regia* genome for 76 first backcross walnuts based on 100 randomly amplified polymorphic DNA (RAPD) markers and restriction fragment-length polymorphisms (RFLPs).

MD, the traits reflected the expected genomic mixing of *J. hindsii*, *J. regia*, their F₁ hybrids and the backcross populations. The differences in the coefficient of variation (C) of the MD for the parental populations (Table 4) suggests that the component traits of the MD may better represent the variability within *J. regia* than *J. hindsii*. If so, this might be an advantage to a breeding program with the goal of selecting those BC₁s morphologically most similar to *J. regia* for further backcrossing and eliminating as much of the *J. hindsii* genome as possible. Alternatively, one might attempt to capture the genetic variability introduced by *J. hindsii* by identifying backcrosses that are morphologically more like *J. hindsii* but with suitable horticultural qualities. Since in this study the MD and PRPG did not correspond, a simple, rapid morphological selection among the backcrosses does not seem promising. Several of the traits used in this analysis were strongly phenotypically correlated (Woeste, 1995). Evaluating a larger number of genetically uncorrelated traits might have produced a better agreement between the MD and PRPG, but the work involved in identifying these traits, scoring the populations and analyzing the data is considerable.

The relationship between the PRPG and the breeding value of the backcrosses is unclear. Selection against *J. hindsii* genome may not be necessary if traits such as shell thickness are strongly affected by only a few genes. In fact, promising variability for nut quality emerged early in the backcrossing (Woeste, 1998). Specific associations between markers and traits of interest might be useful, but because the *J. regia* × *J. hindsii* cross is male sterile and F₁ and BC₁ plants often have very low fertility, these associations could be due to linkage disequilibrium and they would not be present in other crosses. Adding more markers would improve our ability to sample the hybrid genome only marginally since any additional markers would be linked to and in linkage disequilibrium with markers that already exist.

The low correlation between statistical and genetic distance in this walnut BC₁ population is similar to the results of studies by Smith and Smith (1989b) and Beer et al. (1993). In these and other studies involving lines of field crops, researchers typically used large numbers (>10) of agronomic characters as the basis of their morphological evaluation. Despite the large data sets that have been generated in this fashion, the correlation between MD and genetic markers is typically low, in the range of $r = 0.20$ to 0.30 . The results from (sometimes distantly) related populations of field crops are true of backcrosses as well. Selection during backcrossing for recurrent parent genome based on morphology would be an unreliable approach in this walnut BC₁ population. These results also correspond with those from ecological studies. Typically

Table 5. Rank correlations^a among five morphological traits,^b percent *Juglans regia* genome, percent *Juglans regia* genome based on a scaffold set of markers, and Mahalanobis' distance (MD) for first backcross walnuts.

Correlation	Percent <i>Juglans regia</i> genome ^c	Percent <i>Juglans regia</i> genome (scaffold)	MD
SL/SW	-0.23*	-0.22	0.21*
RL/N	0.21	0.06	-0.52***
BL/BW	-0.07	-0.19	-0.19
SW/RL	0.31*	0.24*	-0.62***
PL/RL	0.22	0.07	-0.63***
Percent <i>Juglans regia</i> genome		0.72***	0.23*
Percent <i>Juglans regia</i> genome (scaffold)			0.14

^an ≥ 50 for all correlations.

^bSee text for a description of the traits.

^cEquals PRPG as described in text.

*,***Significant at $P \leq 0.05$ or 0.0001 , respectively.

ecologists consider morphological intermediacy to be reliable evidence of hybridization, but molecular data is often considered to be a more sensitive measure of hybridity, especially when the genetic relationships among populations are complex (Dowling et al., 1989).

In general, when morphological and genetic distance measures are used to evaluate the same population, the results, while sometimes congruent (Atchley et al., 1982b; Villani et al., 1992), are often surprisingly discordant (Atchley et al., 1988; Smith and Smith, 1989b; Lewontin, 1984, 1986) despite the well-accepted theoretical underpinnings of each method. This disparity accentuates the insight that morphological differentiation of populations and individuals is complex; not determined in a simply additive fashion as the sum of genes from each parent or parental population. This was true of the traits considered here, even though they were specifically chosen for this characteristic.

Literature Cited

- Anderson, E. 1949. Introgressive hybridization. Wiley, New York.
- Atchley, W.R., E.V. Nordheim, F.C. Gunsett, and P.L. Crump. 1982a. Geometric and probabilistic aspects of statistical distance functions. *Syst. Zool.* 31:445–460.
- Atchley, W.R., J.J. Rutledge, and D. Crowley. 1982b. A multivariate analysis of direct and correlated response to selection in the rat. *Evolution* 36:677–698.
- Atchley, W.R., S. Newman, and D.E. Cowley. 1988. Genetic divergence in mandible form in relation to molecular divergence in inbred mouse strains. *Genetics* 120:239–253.
- Baum, B.R. and L.G. Bailey. 1984. Taxonomic studies in wall barley (*Hordeum murinum sensu lato*) and sea barley (*Hordeum marinum sensu lato*). II. Multivariate morphometrics. *Can. J. Bot.* 62:2754–2764.
- Becker, W. 1985. Manual of quantitative genetics. 4th ed. Academic Enterprises, Pullman, Wash.
- Beer, S.C., J. Goffreda, T.D. Phillips, J.P. Murphy, and M.E. Sorrells. 1993. Assessment of genetic variation in *Avena sterilis* using morphological traits, isozymes, and RFLPs. *Crop Sci.* 33:1386–1393.
- Brochman, C. 1987. Evaluation of some methods for hybrid analysis, exemplified by hybridization in *Argyranthemum* (Asteraceae). *Nord. J. Bot.* 7:609–630.
- Cheverud, J.M. 1988. A comparison of genetic and phenotypic correlations. *Evolution* 42:958–968.
- Cowen, N.M. and K.J. Frey. 1987. Relationship between three measures of genomic distance and breeding behavior in oats (*Avena sativa* L.) *Genome* 29:97–106.
- Dowling, T.E., G.R. Smith, and W.M. Brown. 1989. Reproductive isolation and introgression between *Notropis cornutus* and *Notropis chrysocephalus* (Family Cyprinidae): Comparison of morphology, allozymes, and mitochondrial DNA. *Evolution* 43:620–634.
- Dudley, J.W. 1994. Comparison of genetic distance estimators using molecular data, p. 3–7. In: Analysis of molecular marker data. Proc. Joint Plant Breeding Symp. Ser., Amer. Soc. Hort. Sci.–Crop Sci. Soc. Amer., Corvallis, Ore., 5–6 Aug. 1994.
- Falconer, D.S. 1981. Introduction to quantitative genetics. 2nd ed. Longman Scientific and Technical, Essex, U.K.
- Fernandez, G.C.J. 1992. Residual analysis and data transformations: Important tools in statistical analysis. *HortScience* 27:297–300.
- Fernando, R.L. and D. Gianola. 1988. Estimation of heritability by offspring–parent regression when observations do not have a common mean. *Theor. Appl. Genet.* 75:803–806.
- Fisher, R.A. 1936. The use of multiple measurements in taxonomic problems. *Ann. Eugenics* 7:179–188.
- Fjellstrom, R.G. 1993. Genetic diversity of walnut (*Juglans* L.) species determined by restriction fragment length polymorphisms, PhD thesis. Univ. of California, Davis.
- Fjellstrom, R.G. and D.E. Parfitt. 1994. RFLP inheritance and linkage in walnut. *Theor. Appl. Genet.* 89:665–670.
- Goodman, M.M. 1968. The races of maize: II. Use of multivariate analysis of variance to measure morphological similarity. *Crop Sci.* 8:693–698.
- Hansche, P.E., V. Beres, and H.I. Forde. 1972. Estimates of quantitative genetic properties of walnut and their implications for cultivar improvement. *J. Amer. Soc. Hort. Sci.* 97:279–285.
- Hillel, J., T. Schaap, A. Haverfeld, A.J. Jeffreys, Y. Plotzky, A. Cahaner, and U. Lavi. 1990. DNA fingerprint applied to gene introgression breeding programs. *Genetics* 124:783–789.
- Hospital, F., C. Chevalet, and P. Mulsant. 1992. Using markers in gene introgression breeding programs. *Genetics* 132:1199–1210.
- Ietswaart, J.H. and A.E. Feij. 1989. A multivariate analysis of introgression between *Quercus robur* and *Q. petraea* in The Netherlands. *Acta Bot. Neerl.* 38:313–325.
- Kesseli, R., O. Ochoa, and R.W. Michelmore. 1991. Variation at RFLP loci in *Lactuca* spp. and origin of cultivated Lettuce (*L. sativa*). *Genome* 34:430–436.
- Lewontin, R.C. 1984. Detecting population differences in quantitative characters as opposed to gene frequencies. *Amer. Nat.* 123:115–124.
- Lewontin, R.C. 1986. A comment on the comments of Rogers and Felsenstein. *Amer. Nat.* 127:733–734.
- Mahalanobis, P.C. 1936. On the generalized distance in statistics. *Proc. Natl. Inst. Sci. India* 2:49–55.
- Moser, H., and M. Lee. 1994. RFLP variation and genealogical distance, multivariate distance, heterosis, and genetic variance in oats. *Theor. Appl. Genet.* 87:947–956.
- Namkoong, G. 1966. Statistical analysis of introgression. *Biometrics* 22:947–956.
- Openshaw, S.J., S.G. Jarboe, and W.D. Beavis. 1994. Marker-assisted selection in backcross breeding, p. 41–43. In: Analysis of molecular marker data. Proc. Joint Plant Breeding Symp. Ser., Amer. Soc. Hort. Sci.–Crop Sci. Soc. Amer., Corvallis, Ore., 5–6 Aug. 1994.
- Pereira-Lorenzo, S., J. Fernandez-Lopez, and J. Moreno-Gonzalez. 1996. Variability and grouping of northwestern Spanish chestnut cultivars. I. Morphological traits. *J. Amer. Soc. Hort. Sci.* 132:183–189.
- Ramos, D. 1985. Walnut orchard management. Regents Univ. of California, Div. Agr. Natural Resources, Oakland, Publ. 21410.
- Smith, J.S.C. and O.S. Smith. 1989a. The description and assessment of distances between inbred lines of maize: I. The use of morphological traits as descriptors. *Maydica* 34:141–150.
- Smith, J.S.C. and O.S. Smith. 1989b. The description and assessment of distances between inbred lines of maize: II. The utility of morphological, biochemical and genetic descriptors and a scheme for testing of distinctiveness between inbred lines. *Maydica* 34:151–161.
- Souza, E., and M.E. Sorrells. 1991. Relationships among 70 North American oat germplasm: I. Cluster analysis using quantitative characters. *Crop Sci.* 31:599–605.
- Tanksley, S.D., H. Medina-Filho, and C.M. Rick. 1981. The effect of isozyme selection on metric characters in an interspecific backcross of tomato—Basis of an early screening procedure. *Theor. Appl. Genet.* 60:291–296.
- University of California Statewide Integrated Pest Management Project. 1982. Integrated pest management for walnuts. *Agr. Sci. Publ. Div. Agr. Sci., Berkeley, Publ.* 3270.
- Villani, F., M. Pigliucci, M. Lauteri, M. Cherubini, and O. Sun. 1992. Congruence between genetic, morphometric, and physiological data on differentiation of Turkish chestnut (*Castanea sativa*). *Genome* 35:251–256.
- Williams, C.E. and D.A. St. Clair. 1993. Phenetic relationships and levels of variability detected by restriction fragment length polymorphism and random amplified polymorphic DNA analysis of cultivated and wild accessions of *Lycopersicon esculentum*. *Genome* 36:619–630.
- Woeste, K.E. 1995. Genetic markers for disease resistance breeding in Persian walnut (*Juglans regia* L.), PhD thesis. Univ. of California, Davis.
- Woeste, K., G.H. McGranahan, and R. Bernatzky. 1996. Randomly amplified polymorphic DNA loci from a walnut backcross [*Juglans hindsii* x *J. regia*] x *J. regia*. *J. Amer. Soc. Hort. Sci.* 121:358–361.
- Woeste, K., C.A. Leslie, and G.H. McGranahan. 1998. Backcross breeding walnuts for resistance to the cherry leafroll virus. *Acta Hort.* (In press.)
- Yang, S.Y. and R.K. Selander. 1968. Hybridization in the grackle *Quiscalus quiscula* in Louisiana. *Syst. Zool.* 17:107–143.
- Zeyl, C.W. and L.A. Lowcock. 1989. A morphometric analysis of hybrid salamanders (genus *Ambystoma*) and their progenitors on Kelly's Island in Lake Erie. *Can. J. Zool.* 67:2376–2383.