SHORT COMMUNICATION

Genetic divergence is not the same as phenotypic divergence

Marcin Kozak · Jan Bocianowski · Alina Liersch ·
Małgorzata Tartanus · Iwona Bartkowiak-Broda ·
Fernando A. Piotto · Ricardo A. Azevedo

Abstract Far too often, phenotypic divergence has been misinterpreted as genetic divergence, and based on phenotypic divergence, genetic divergence has been indicated. We have attempted to disprove this statement and call for the differentiation of phenotypic and genotypic variation.

Keywords Genotype · Phenotype · Plant breeding

M. Kozak
Department of Experimental Design and Bioinformatics,
Warsaw University of Life Sciences, Warsaw, Poland

J. Bocianowski (✉)
Department of Mathematical and Statistical Methods,
Poznań University of Life Sciences, Wojska Polskiego 28,
60-637 Poznań, Poland
e-mail: jboc@up.poznan.pl

A. Liersch · I. Bartkowiak-Broda
Department of Oilseed Crops, Plant Breeding and
Acclimatization Institute, National Research Institute,
Strzeszyńska 36, 60-479 Poznań, Poland

M. Tartanus
Research Institute of Pomology and Floriculture,
Pomologiczna 18, 96-100 Skierniewice, Poland

F. A. Piotto · R. A. Azevedo
Departamento de Genética, Escola Superior de
Agricultura Luiz de Queiroz, Universidade de São Paulo,
Piracicaba CEP 13418-900, Brazil

Genetic divergence has been a hot topic in plant breeding in recent decades, and still is. It should be understood as the divergence of the gene pool of a population from the gene pools of other populations, which can be due to mutation, genetic drift and selection. Hence genetic divergence describes genetic divergence within the population, and as such is not the same as phenotypic divergence (or diversity, as it is usually termed—hereafter we will only use the term divergence), which describes divergence within the population in terms of phenotypic traits.

Unfortunately, all too often, genetic divergence is misunderstood as phenotypic divergence (Anjani 2005; Arriel et al. 2007; Bisht et al. 2007; Bose and Pradhan 2005; Debnath et al. 2008; Gashaw et al. 2007; Goswami et al. 2006; Hossain 2006; Kabir et al. 2009; Khan et al. 2008; Kumar 2008; Shanmuganathan et al. 2006; Singh et al. 2006; Thayumanavan et al. 2009). This approach can be found in interpreting experiments with a pool of genotypes being studied in terms of various plant and crop traits, and a simplification is made that if two genotypes are similar in such a multivariate phenotypic way, they are genetically similar. Therefore this line of thinking suggests that if members of a pool of genotypes are diverse in terms of many traits, they are probably genetically divergent over much of the genome.

The idea of associating phenotypes with genotypes is somewhat basic and common in plant breeding, as well as in the history of plant domestication carried out through the selection of better plants for

agricultural use. The recent widespread use of molecular markers has been helpful in establishing the direct association between phenotype and genotype. This has been carried out for several reasons, including mapping genes and quantitative trait loci (QTL), and the evaluation of divergence and genetic distance among species accessions and their relatives. However, several reports have shown that the correlation between morphological traits and genotypes is relatively low (Bar-Hen et al. 1995; Bernet et al. 2003; Kwon et al. 2005; Lefebvre et al. 2001; Tommasini et al. 2003). In addition, some morphological traits, for example maize flowering time (Buckler et al. 2009), have been shown to be controlled by many loci. Different morphological traits might therefore have different levels of indication about genetic divergence. Compared with morphological traits controlled by single major genes, the phenotypic divergence computed from complex morphological traits controlled by many genes would therefore better reflect genetic divergence.

The problem is that what here constitutes a basis for such an interpretation of genetic divergence is actually phenotypic divergence within the pool of genotypes. Is there any evidence that large phenotypic divergence is equivalent to large genetic divergence? It is well established that the phenotypes of a given genotype vary from environment to environment, and from replication to replication in one environment. Furthermore, in some cases the high divergence of phenotypes is intrinsic to the species. Tomato is a classical example of a strong contrast between phenotypic and genetic variability, since this plant species exhibits considerable morphological divergence whilst its genetic divergence is reduced, and it is estimated that tomato only has around 5% of the variation shown by related species (Miller and Tanksley 1990). Even the major domestication characteristic of tomato (increase in fruit size) is due to mutations in about six QTL (Tanksley 2004), in particular QTL *fw2.2*, which by itself is responsible for about 30% of the fruit size phenotype (Frary et al. 2000). Interestingly, although the domestication process for this species has been severe, leading to a narrow genetic base, it is still possible to find several morphological markers for almost all characteristics (e.g. shoots, leaves, fruits, flowers). The divergence in morphological markers is so large that it was possible to produce linkage maps

(Rick and Yoder 1988) which were and are still used by geneticists and plant breeders around the world, associated with maps of modern molecular markers (www.solgenomics.net). Although it is possible to identify a large number of morphological markers in tomato, the related genes represent a small percentage of the total gene pool of commercial cultivars. As an example, it is possible to cite some classical markers such as *potato leaf (c)*, *Beta (B)*, *white flower (wf)*, *yellow flesh (r)*, *anthocyaninless (a)*, *hairless (h)*, among others (www.tgrc.ucdavis.edu).

On the other hand, it is also possible to identify cases where the converse occurs, e.g. when plants express considerable morphological similarity, but are genetically very different. A typical example is the Chalco race of teosinte (*Zea mays* ssp. *mexicana*), which is a maize-mimetic weed (Wilkes 1967) found in maize fields of Mexico. This weed looks very much like a maize plant up to the stage of flowering; however, the isoenzyme profiles of Chalco teosinte and maize plants are genetically divergent (Doebley et al. 1987). This suggests that Chalco teosinte can adopt a phenotypic disguise from convergent evolution of the phenotype to avoid being weeded out, whilst still retaining its genetic difference. Therefore, in the same way it cannot be said that phenotypic distance can be used as a replacement for measuring genetic distance, or that genetic distance can be used as a replacement for measuring phenotypic distance.

Therefore, the choice of traits to include in the determination of phenotypic divergence also counts, and for various sets the same pool of genotypes will differ in this divergence. This means that phenotypic divergence should be analyzed and discussed in a particular context, given by traits constituting the basis for a coefficient describing the divergence.

Let us consider an example from an actual experiment. Eighteen parental lines of F_1 hybrids of winter oilseed rape based on the CMS *ogura* hybridization system were examined. The genetic polymorphisms of the lines were analyzed using 597 random amplified polymorphic DNA, amplified fragment length polymorphism and isozyme markers. In addition, the lines were evaluated in a field experiment in two locations, in two years (2002/2003 and 2003/2004), where the four combinations of location and year were treated as environments. The experiments were laid out in a randomized complete block design with four replications, conducted at the

Experimental Station of Wielichowo in Zielęcin (52°10′N, 16°23′E) and Plant Breeding Company Strzelce Ltd in Borowo (52°07′N, 16°46′E), Poland. The following phenotypic traits were of interest: seed yield, pod length, number of seeds per pod, thousand seed weight, beginning of flowering, length of flowering, content of oil, palmitic acid, stearic acid, oleic acid, linoleic acid, linolenic acid, eicosenoic acid, and alkenyl glucosinolates (gluconapine, gluco-brassicanapine, progoitrine, 4-hydroxybrassicine, total alkenyl glucosinolates and total glucosinolates).

The Mahalanobis distances (Mahalanobis 1936) between the lines, determined using these traits, can be treated as phenotypic similarities between the lines, established separately for the four environments. The phenotypic divergence can be interpreted in the context of the above-mentioned phenotypic traits. The genetic similarity between the pairs of lines was determined using the Nei and Li (1979) coefficient. Thus, these two distances can be thought of as indices of phenotypic and genetic divergence, respectively. In Fig. 1, it can be seen that they do not have to be related at all: genetic divergence was associated with phenotypic divergence only in Zielęcin 2004, while for the three other environments no such association was detected. It can also be seen that the phenotypic divergence varies from environment to environment, and that even relationships among phenotypic distances in two environments do not have to be linear and strong.

The basis of morphological and genotypic divergences or similarities could be considered as a complex genetic system. Differences among varieties of the same or distinct species may be due to allelic variation and differential gene expression associated with morphological and non-physiological traits, in addition to the genotype-by-environment interaction of each individual locus, which can create much phenotypic divergence and noise. Phenotypic divergence reflects a small fraction (often unknown) of the genes and environmental interaction, which are in general strongly affected by the environment. However, genetic divergence evaluated with molecular markers provides a more precise and potentially more representative portrayal of divergence for the genome as a whole, but this will depend upon genomic coverage.

So, why should the two terms be mixed up and phenotypic divergence be treated as genetic divergence? The bottom line is that the latter can—and usually does—affect the former, but so does the environment. These two terms are not equivalent, so it would be beneficial to stop equating phenotypic and genotypic divergence as always representing the same measure of divergence.

Fig. 1 Associations among genetic similarities of *B. napus* L. lines and phenotypic (Phen) distances determined in four environments (*B3*, *B4*: Borowo 2003 and 2004; *Z3*, *Z4*: Zielęcin 2003 and 2004, Poland). A locally weighted regression line is superimposed over the panels to help grasp the relationship between the corresponding variables

References

Anjani K (2005) Purple-coloured castor (*Ricinus communis* L.)—a rare multiple resistant morphotype. Curr Sci 88(2):215–216

Arriel NHC, Mauro AOD, Arriel EF, Unêda-Trevisoli SH, Costa MM, Bárbaro IM, Muniz FRS (2007) Genetic divergence in sesame based on morphological and agronomic traits. Crop Breed Appl Biotechnol 7:253–261
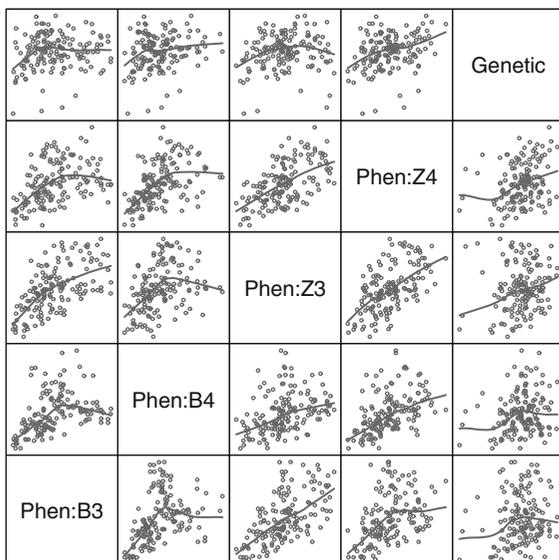
Bar-Hen A, Charcosset A, Bourgoin M, Guiard J (1995) Relationship between genetic markers and morphological traits in a maize inbred lines collection. Euphytica 84:145–154

Bernet GP, Bramardi S, Calvache D, Carbonell EA, Asins MJ (2003) Applicability of molecular markers in the context of protection of new varieties of cucumber. Plant Breed 122:146–152

Bisht V, Singh B, Rao KS, Maikhuri RK, Nautiyal AR (2007) Genetic divergence of paddy landraces in Nanakosi micro-watershed of Uttarakhand Himalaya. J Trop Agric 45(1–2):48–50

Bose LK, Pradhan SK (2005) Genetic divergence in deepwater rice genotypes. J Cent Eur Agric 6(4):635–640

Buckler ES, Holland JB, Bradbury PJ, Acharya CB, Brown PJ, Browne C, Ersoz E, Flint-Garcia S, Garcia A, Glaubitz JC, Goodman MM, Harjes C, Guill K, Kroon DE, Larsson S, Lepak NK, Li H, Mitchell SE, Pressoir G, Peiffer JA, Rosas MO, Rocheford TR, Romay MC, Romero S, Salvo S, Villeda HS, Sofia da Silva H, Sun Q, Tian F, Upadyayula N, Ware D, Yates H, Yu J, Zhang Z, Kresovich S, McMullen MD (2009) The genetic architecture of maize flowering time. Science 325:714–718

Debnath NR, Rasul MG, Sarker MMH, Rahman MH, Paul AK (2008) Genetic divergence in buckwheat (Fagopyrum esculentum Moench.). Int J Sustain Crop Prod 3(2):60–68

Doebley J, Goodman MM, Stuber CW (1987) Patterns of isozyme variation between maize and Mexican annual teosinte. Econ Bot 41(2):234–246

Frary A, Nesbitt TC, Frary A, Grandillo S, Van-Der-Knaap E, Cong B, Liu J, Meller J, Elber R, Alpert KB, Tanksley SD (2000) A quantitative trait locus key to the evolution of tomato fruit size. Science, New York 289:85–88

Gashaw A, Mohammed H, Singh H (2007) Genetic di-vergence in selected durum wheat genotypes of ethiopian plasm. Afr Crop Sci J 15(2):67–72

Goswami PK, Baruah PP, Behl RK (2006) Genetic divergence in relation to combining ability and transgression in Indian mustard (Brassica juncea L.). World J Agric Sci 2(4):390–393

Hossain MA (2006) Screening of mungbean (Vigna radiata) genotypes for nutrient stress tolerance. J Agric Rural Dev 4(1–2):129–135

Kabir MY, Khan ASMMR, Hassain MS (2009) Genetic divergence in pointed gourd. J Agric Rural Dev 7(1&2):87–92

Khan ASMR, Rabbani MG, Siddique MA, Hossain I (2008) Study on genetic diversity of pointed gourd using morphological characters. Bangladesh J Agr Res 33(3):607–616

Kumar BMD (2008) Genetic divergence in red rice. Karnataka J Agric Sci 21(3):346–348

Kwon YS, Lee JM, Yi GB, Yi SI, Kim KM, Soh EH, Bae KM, Park EK, Song IH, Kim BD (2005) Use of SSR markers to complement tests of distinctiveness, uniformity, and stability (DUS) of pepper (Capsicum annuum L.) varieties. Mol Cell 19(3):428–435

Lefebvre V, Goffinet B, Chauvet JC, Caromel B, Signoret P, Brand R, Palloix A (2001) Evaluation of genetic distances between pepper inbred lines for cultivar protection purposes: comparison of AFLP, RAPD and phenotypic data. Theor Appl Genet 102:741–775

Mahalanobis PC (1936) On the generalized distance in statistics. Proc Natl Inst Sci India 12:49–55

Miller JC, Tanksley SD (1990) RFLP analysis of phylogenetic relationships and genetic variation in the genus Lycopersicon. Theor Appl Genet 80:437–448

Nei M, Li W (1979) Mathematical model for studying genetic variation in terms of restriction endonucleases. Proc Natl Acad Sci USA 76:5269–5273

Rick CM, Yoder JI (1988) Classical and molecular genetics of tomato: highlights and perspectives. Annu Rev Genet 22:281–300

Shanmuganathan M, Gopalan A, Mohanra K (2006) Genetic variability and multivariate analysis in pearl millet (Pennisetum glaucum (L.) R. Br.) germplasm for dual purpose. J Agric Sci 2(1):73–80

Singh PK, Mishra MN, Hore DK, Verma MR (2006) Genetic divergence in lowland rice of north eastern region of India. Commun Biometry Crop Sci 1(1):35–40

Tanksley SD (2004) The genetic, developmental and molecular bases of fruit size and shape variation in tomato. Plant Cell 16:S181–S189

Thayumanavan S, Kannapiran S, Annamalai A (2009) Genetic divergence analysis for certain yield and quality traits in rice (Oryza sativa L.) grown in irrigated saline low land of Annamalainagar, South India. J Cent Eur Agric 10(4):405–410

Tommasini L, Batley J, Arnold GM, Cooke RJ, Donini P, Lee D, Law JR, Lowe C, Moule C, Trick M, Edwards KJ (2003) The development of multiplex simple sequence repeat (SSR) markers to complement distinctness, uniformity and stability testing of rape (Brassica napus L.) varieties. Theor Appl Genet 106:1091–1101

Wilkes HG (1967) Teosinte: the closest relative of maize. Bussey Institution, Harvard University, Cambridge, Massachusetts, p 159