EPIGENETIC INFORMATION – UNEXPLORED SOURCE OF NATURAL VARIATION

Ovidiu Paun* & Mark W. Chase

Jodrell Laboratory, Royal Botanic Gardens, Kew, Richmond, Surrey TW9 3DS, U.K.
* Corresponding author: ovidiu.paun@univie.ac.at

Abstract. Rapid progress is being made at the population-level in orchids, with a series of new molecular techniques being applied. One of the major problems observed in several groups of temperate orchids has been that ecologically distinct "taxa" do not appear to be genetically distinct. For example, we know that *Dactylorhiza traunsteineri*, *D. majalis,* and *D. ebudensis* are the products of hybridization between *D. fuchsii* and *D. incarnata*, but they have different ecologies and distributions within northwestern Europe. By comparing fingerprinting analyses of expressed regions to fingerprinting studies of methylation-sensitive sites in genomic DNA, we can detect patterns that indicate that some of these differences are due to changing epigenetic effects, which have been shown in several groups to be subject to environmental influence. Thus, taxa that are ecologically distinct but still appear genetically uniform may be the result of altered epigenetic controls of gene expression without any change in the underlying genetic material.

Resumen. Con la aplicación de nuevas técnicas moleculares, rápidos avances se han hecho para comprender la organización de las orquídeas a nivel de poblaciones. Uno de los mayores problemas encontrados en varios grupos de orquídeas de zonas temperadas, ha sido que los taxa ecológicamente diferentes, al parecer son genéticamente similares. Por ejemplo, sabemos que *Dactylorhiza traunsteineri*, *D. majalis*, y *D. ebudensis* son resultado de la hibridización de *D. fuchsii* y *D. incarnata*, pero éstas tienen ecología y distribución diferentes dentro del nor-oeste de Europa. A través de la comparación de análisis "fingerprint" de regiones expresadas, con estudios "fingerprint" de sitios sensibles a la metilación en ADN genómico, podemos detectar patrones que indican que algunas de estas diferencias se deben a efectos epigenéticos, los cuales han sido encontrados en varios grupos que han sido expuestos a la influencia del ambiente. Así, taxa que son ecológicamente diferentes, pero que son al parecer genéticamente uniformes, podrían ser el resultado de la expression de genes, debido a una alteración de los controles epigenéticos, pero sin que haya ocurrido ningún cambio en el material genético.

Key words: orchids, hybridization, fingerprinting analyses, epigeny, environmental influence

DNA sequences of all cells in an organism are essentially identical, but their transcription program is specific, depending on the cell type, developmental stage, age, location, etc. This critical feature of multicellular organisms is achieved through an array of epigenetic signals that are not encoded in the primary DNA sequence. Epigenetic signposts include DNA methylation (addition of a methyl group to a cytosine; Martienssen and Colot 2001), modifications of histones (core proteins that package DNA into chromosomal particles; Grewal & Moazed 2003), and small regulatory RNAs (Großhans & Filipowicz 2008). These signals control the activity of genes and transposable elements in a relatively flexible manner (Rapp & Wendel 2005) by providing differential access to underlying genetic information to modulate the identity of cells.

In addition to the extensive involvement in developmental control and parent-of origin imprinted gene expression (Steimer *et al.* 2004, Henderson & Jacobsen 2007), epigenetic variants (epialleles) also control the organismic response to environmental conditions. Surprisingly, recent studies in various model organisms – including monozygotic twin humans (Fraga *et al.* 2005) – have indicated that some epialleles can be inherited

across generations through mechanisms that are not based on the primary DNA sequence. Accumulating evidence indicates that similarly to controlling cell differentiation within an organism, epigenetics can contribute, even in the absence of genetic or environmental heterogeneity, an additional layer of information that influences heritable phenotypic variation between individuals and evolutionary processes that act on this variation.

This may be particularly true for plants, where redundant copies of many genes are spread across the genomes as a result of ancient whole-genome duplication events (i.e., paleopolyploidization; Soltis *et al.* 2009) or individual gene duplication. A duplicated genetic background allows for increased levels of epigenetic variation that sorts out possible genetic variants. Moreover, epigenetic response mechanisms may be particularly important for immobile organisms, such as plants, that generally cannot respond within the same generation to environmental stimuli via relocation or behavioral alterations. Indeed, many examples of spontaneous heritable epialleles that demonstrably affect key phenotypic characters, such as pigmentation, floral shape, and pathogen resistance, have been discovered in flowering plants (Jablonka & Raz 2009), reflecting also the characteristically late partitioning of reproductive and vegetative cell lineages in higher plants.

Therefore, epigenetics could significantly improve our understanding of the natural mechanisms underlying phenotypic variation and the response of organisms to environmental change (Richards 2006, Bossdorf *et al.* 2008). However, epigenetic markers are at present only starting to be investigated in relation to evolutionary questions. We have little understanding regarding even basic questions, such as the extent and structure of epigenetic variation within and among natural populations. Recently, epigenetic markers have started to be used to investigate evolutionary questions related to stabilization of recently formed polyploids and evolutionary success. Several studies have provided evidence that epigenetic mechanisms activate dormant mobile elements and silence redundant genes immediately after polyploidization events, providing a flexible and reversible way of stabilizing cellular processes (see Paun *et al.* 2007, for a

review). However, little information is available on the stability of these epigenetic changes and their potential long-term evolutionary implications. We have studied a group of closely related polyploid species of orchids that developed during the last part of the Quaternary and found convincing evidence that epigenetic, rather than genetic, differentiation has been important for their diversification, driving ecological allopatry and reproductive isolation.

The case of *Dactylorhiza* allopolyploids

Dactylorhiza, a genus of temperate terrestrial orchids, has extensively evolved in a reticulate fashion resulting in significant but often subtle morphological and ecological variation that challenges species delimitation. Several *Dactylorhiza* allotetraploid ($2n = 80$) species have been repeatedly formed by hybridization between two broadly defined and geographically widespread parental lineages: the diploid ($2n = 40$) marsh-orchid, *D. incarnata* (L.) Soó *s.l.*, and the spotted orchid, *D. maculata* (L.) Soó *s.l.* (including, among others, diploid *D. fuchsii* (Druce) Soó; Heslop-Harrison 1968, Hedrén 1996, Hedrén *et al.* 2001, Pillon *et al.*, 2007, Hedrén *et al.* 2008). As a result, the taxonomy of many dactylorchids is widely considered complicated, causing long-standing controversies. The different sibling polyploid taxa can occur sympatrically, but they have different overall ecological requirements and distinct distributions. The full significance of the highly iterative polyploid evolution in this complex is not yet understood. It is intriguing that allotetraploid derivatives of the same progenitor pair can exhibit contrasting morphological and ecological properties that are maintained in spite of partly overlapping distributions.

Examples of such cases include European allotetraploids *D. majalis* (Rchb.) P.F.Hunt & Summerh. *s.str.*, *D. traunsteineri* (Saut. ex Rchb.) Soó *s.l.* (Fig. 1) and *D. ebudensis* (Wief. ex R.M.Bateman & Denholm) P.Delforge, each derived multiple times (except the last, a localized endemic, which has most probably a single origin) from unidirectional hybridization between diploids *D. fuchsii* (in all cases the maternal parent) and *D. incarnata* (Pillon *et al.*, 2007, and references therein). Combining the degree of concerted evolution in

FIGURE 1. The allotetraploid *Dactylorhiza traunsteineri* at a natural site in Yorkshire, Britain (Photo: O. Paun).

ITS alleles, and in agreement with the patterns of morphology and ecological preference, *D. majalis* is more derived and genetically homogeneous, inferred to be the oldest of the three allotetraploids and to have passed through glacially induced bottlenecks in southern Eurasia. It has a fairly wide ecological tolerance of soil moisture and occurs at present in damp meadows and fens in western and central Europe, the Baltic region, and northern Russia. In contrast, *D. traunsteineri* is a more recently evolved set of allotetraploids that is more heterogeneous and still maintains both parental ITS alleles (Pillon *et al.*, 2007). It probably originated post-glacially and at present shows a more localized and disjunct distribution in northwestern and central Europe (i.e., Britain, Scandinavia, and the Alps). It has narrow tolerances of both soil moisture and pH and grows in calcareous fens and marshes. A third allotetraploid, *D. ebudensis*, is a narrow endemic (at present forming a single population) in northwestern Scotland and may be as young as or younger than *D. traunsteineri*. The coastal dune habitat occupied

by *D. ebudensis* indicates its relatively narrow tolerances of both soil moisture and pH.

Despite ecological and morphological distinctiveness, neutral genetic differentiation between sibling *Dactylorhiza* allopolyploids, as studied with various molecular markers, including allozymes, ITS sequences, AFLPs, and plastid and nuclear microsatellites, has proved to be rarely detectable (Hedrén *et al.*, 2001; Pillon *et al.*, 2007). By contrast, a fingerprinting analyses of gene expression patterns (cDNA-AFLP; Paun *et al.*, 2007) resulted in a geographic structure (Paun *et al.*, 2010), grouping samples according to their provenance (i.e., Britain, Scandinavia, Alps, and Pyrenees). This may reflect the multiple independent origins of each tetraploid taxon (except *D. ebudensis*) or may provide evidence of local adaptation and/or stronger regional gene flow. However, several expression patterns have been found to be species-specific and clearly discriminate between the allopolyploid species (Paun *et al.*, 2010).

Moreover, epigenetic variation at loci spread across the genome, which was studied using methylation-sensitive amplified polymorphism (MSAP), clearly separates the three allopolyploids studied here (Fig. 2), in stark contrast to patterns of genome-wide genetic data (Hedrén *et al.*, 2001). The MSAP technique is similar to standard AFLP (Vos *et al.*, 1995) but uses two methylation-sensitive restriction isoschizomers (e.g., MspI and HpaII) as frequent cutters in parallel batches (Baurens *et al.*, 2003). The two isoschizomers recognize the same DNA sequence (5`-CCGG) but differ in their sensitivity to DNA methylation, so that HpaII and MspI do not recognize the restriction site when the internal, and respectively, external cytosine is methylated. Comparisons of the paired profiles for each individual allow precise assessment of methylation status at each restriction site.

Unexpectedly, methylation status analysis at 332 genome-wide MSAP markers resulted in species-specific patterns and confirm previous hypotheses of their evolutionary history (Heslop-Harrison, 1968; Pillon *et al.*, 2007). Indeed, the migration-induced genetic bottleneck triggered within *D. majalis* by the profound change of climatic conditions during the last glaciation seems to have homogenized methylation status among individuals and left only a
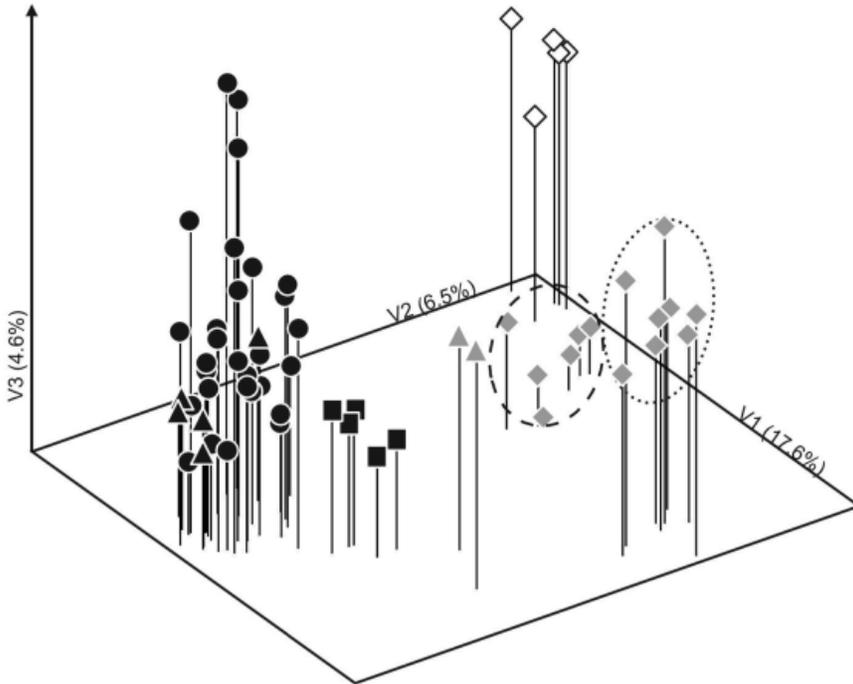
FIGURE 2. Principal coordinates analysis (goodness of fit 0.87 at $P = 0.001$) of methylation status of allotetraploid *D. majalis* (black symbols), *D. traunsteineri* (gray symbols), and *D. ebudensis* (open symbols). Geographical provenance is indicated by symbol shapes: squares, N. Pyrenees; diamonds, Britain; triangles, Scandinavia; circles, E. Alps. The dotted line encloses samples from Yorkshire, England, and the dashed line samples from N.W. Scotland.

weak within-species geographic differentiation (Fig. 2). As result of its postglacial formation and present disjunct distribution, the genome-wide methylation patterns of *D. traunsteineri* are more heterogeneous and correlate more clearly with geography.

Although containing information from both coding and non-coding DNA regions, the methylation markers studied reflect better gene-expression differences in the three allopolyploid siblings than neutral genetic data (Mantel tests, $r = 0.81$ vs. $r = 0.67$ both significant at $P < 0.01$). Therefore, it seems that habitat preference shapes similar expression patterns in some, but not all, of the independent allopolyploidization events in this group, operating via epigenetic regulatory effects under environmental influence rather than via sequence (genetic) divergence (Paun *et al.*, in press).

To test further our hypothesis of adaptation through selection of epigenetic variants, we have performed scans for outlier epigenetic loci, looking for adaptive epiloci that have been shaped by natural selection. We have used two complementary approaches: 1) multiple univariate logistic regressions between epigenetic data and eco-climatic variables, performed using SAM (Joost *et al.*, 2007); and 2) a Bayesian outlier locus approach as implemented in BayeScan (Foll and Gaggiotti, 2008). The latter approach estimates the posterior probability of each locus being under selection, and it is able to differentiate the type of selection each marker was subjected to (divergent vs. purifying). Within our methylation data, BayeScan identified 23 epiloci as being under divergent selection; SAM pinpointed 14 methylation markers as being adaptive, and ten epiloci were identified by both approaches. For most of the outlier epiloci, their presence-absence patterns correlated with water availability, which seems to be a key ecological factor in driving environmental allopatry within the three allopolyploid species. If maintained over evolutionary timescales, environmental allopatry may effectively limit dispersal between populations,

thus promoting divergence via the stochastic effects of drift (Nosil *et al.*, 2009) and/or further selection. Therefore, epigenetic regulatory processes play a clear role in incipient adaptation and evolution of the allopolyploid *Dactylorhiza* species by influencing primary phenotypic diversity at the interface between genetics and the environment. The fact that recent evolutionary history is discernible in epigenetic patterns indicates that such markers could be used for investigations in closely related, rapidly radiating groups when genetic markers may fail to provide relevant information.

LITERATURE CITED

Baurens, F.-C., F. Bonnot, D. Bienvenu, S. Causs & T. Legavre. 2003. Using SDAFLP and MSAP to assess CCGG methylation in the banana genome. Pl. Molec. Biol. Reporter 21: 339-348.

Bossdorf, O., C. L. Richards & M. Pigliucci. 2008. Epigenetics for ecologists. Ecol. Letters 11: 106-115.

Foll, M. & O. Gaggiotti. 2008. A genome-scan method to identify selected loci appropriate for both dominant and codominant markers: a Bayesian perspective. Genetics 180: 977-993.

Fraga, M. F., E. Ballestar, M. F. Paz, S. Ropero, F. Setien, M. L. Ballestar, D. Heine-Suner, J. C. Cigudosa, M. Urioste, J. Benítez, M. Boix-Chornet, A. Sánchez-Aguilera, C. Ling, E. Carlsson, P. Poulsen, A. Vaag, Z. Stephan, T. D. Spector, Y.-Z. Wu, C. Plass & M. Esteller. 2005. Epigenetic differences arise during the lifetime of monozygotic twins. Proc. Nati. Ac. Sc., USA 102: 10604-10609.

Grewal, S. I. S. & D. Moazed. 2003. Heterochromatin and epigenetic control of gene expression. Science 301: 798-802.

Großhans, H. & W. Filipowicz. 2008. The expanding world of small RNAs. Nature 451: 414-416.

Hedrén, M. 1996. Genetic differentiation, polyploidization and hybridization in northern European *Dactylorhiza* (Orchidaceae): evidence from allozymes markers. Pl. Syst. Evol. 201: 31-55.

Hedrén, M., M. F. Fay & M. W. Chase. 2001. Amplified fragment length polymorphisms (AFLP) reveal details of polyploidy evolution in *Dactylorhiza* (Orchidaceae). Amer. J. Bot. 88: 1868-1880.

Hedrén, M., S. Nordström & D. Ståhlberg. 2008. Polyploid evolution and plastid DNA variation in the *Dactylorhiza incarnata/maculata* complex (Orchidaceae) in Scandinavia. Molec. Ecol. 17: 5075-5091.

Henderson, I. R. & S. E. Jacobsen. 2007. Epigenetic inheritance in plants. Nature 447: 418-424.

Heslop-Harrison, J. 1968. Genetic system and ecological habit as factors in dactylorchid variation. Jahresber. Naturwiss. Ver. Wuppertal 21+22: 20-27.

Jablonka, E. & G. Raz. 2009. Transgenerational epigenetic inheritance: prevalence, mechanisms, and implications for the study of heredity and evolution. Quarter. Rev. Biol. 84: 131-176.

Joost, S., A. Bonin, M. W. Bruford, C. Conord, G. Erhardt & P. Taberlet. 2007. A spatial analysis method (SAM) to detect candidate loci for selection: towards a landscape genomics approach to adaptation. Molec. Ecol. 16: 3955-3969.

Martienssen, R. A. & V. Colot. 2001. DNA methylation and epigenetic inheritance in plants and filamentous fungi. Science 293: 1070-1074.

Nosil, P., D. J. Funk& D. Ortíz-Barrientos. 2009. Divergent selection and heterogeneous genomic divergence. Molec. Ecol.18: 375-402.

Paun, O., Bateman, R. M., Fay, M. F., Hedrén, M., Civeyrel, L. and Chase, M. W. In press. Stable epigenetic effects and adaptation in allopolyploid orchids (*Dactylorhiza*: Orchidaceae). *Molecular Biology and Evolution* doi: 10.1093/molbev/msq 150.

Paun, O., Fay, M. F., Soltis, D. E., and Chase, M. W. 2007. Genetic and epigenetic alterations after hybridization and genome doubling. *Taxon* 56: 649-656.

Paun, O., J. A. Luna, M. F. Fay, R. M. Bateman & M. W. Chase. 2010. Genomic responses drive adaptation in allotetraploid species of *Dactylorhiza* (Orchidaceae; Orchidinae). Pp. 169-192 *in*: O. Seberg, G. Petersen, A. S. Barford & J. I. Davis (eds.). Diversity, Phylogeny, and Evolution of the Monocotyledons. Aarhus University Press, Copenhagen, Denmark.

Pillon, Y., M. F. Fay, M. Hedrén, R. M. Bateman, D. S. Devey, A. B. Shipunov, M. van der Bank & M. W. Chase. 2007. Evolution and temporal diversification of western European polyploid species complexes in *Dactylorhiza* (Orchidaceae). *Taxon* 56: 1185-1208.

Rapp, R. A. and Wendel, J. F. 2005. Epigenetics and plant evolution. New Phytol. 168: 81-91.

Richards, E. J. 2006. Inherited epigenetic variation - revisiting soft inheritance. Nature Rev. Genet. 7: 395-402.

Soltis, D. E., V. A. Albert, J. Leebens-Mack, C. D. Bell, A. H. Paterson, C. Zheng, D. Sankoff, C. W. dePamphilis, P. K. Wall & P. S. Soltis. 2009. Polyploidy and angiosperm diversification. Amer. J. Bot. 96: 336-348.

Steimer, A., H. Schöb & U. Grossniklaus. 2004. Epigenetic control of plant development: new layers of complexity. Current Opinion Pl. Biol. 7: 11-19.

Vos, P., R. Hogers, M. Bleeker, M. Reijans, T. van de Lee, M. Hornes, A. Frijters, J. Pot, J. Peleman, M. Kuiper & M. Zabeau. 1995. AFLP: a new technique for DNA fingerprinting. Nucleic Acids Res. 23: 4407-4414.