© NIAB 2010
ISSN 1479-2621
Plant Genetic Resources: Characterization and Utilization (2011) 9(1); 1–3
doi:10.1017/S1479262110000432

Germplasm resources in legumes

T. H. Noel Ellis*†
Department of Crop Genetics, John Innes Centre, Colney Lane, Norwich NR4 7UH, UK

Abstract
The collected articles in this volume are discussed in their wider context.

In this issue of *Plant Genetic Resources*, we have collected articles that describe our current understanding of genetic diversity in *ex situ* collections of legume germplasm. These articles are not a comprehensive description for all legume crop species, but the range covered gives a reasonable flavour of current concerns in both inbreeding and outbreeding species (clover, Abberton and Thomas, and sanfoin, Hayot-Carbonero *et al.*). For soybean (Qui *et al.*, Nelson), chickpea (Upadhyaya *et al.*, Krishnamurthy *et al.*) and lentil (Singh *et al.*, Tulli *et al.*), this is viewed from more than one perspective, while for common bean, Díaz *et al.* consider the structure of germplasm where both the Andean and Mesoamerican domesticates were introduced.

A combination of methodological developments has changed the amount of data available to us for the understanding of genetic diversity. This has in turn changed our view of the ways to approach germplasm collections generally, including ways of generating and of analysing data. Both of these topics are discussed in several of the articles in this volume. The article on pea germplasm by Smykal *et al.*, for example, includes a comparison of the results from different Bayesian modelling methods.

Developments in molecular taxonomy have come to maturity since the seminal paper by Chase *et al.* (1993) first attempted a comprehensive description of angiosperm molecular phylogeny. The promise of that study has come good, and for legumes, a detailed phylogeny with many divergence times between groups has been presented by Lavin *et al.* (2005). Some of these dates were established with reference to the fossil record, while others are interpolated from patterns of sequence divergence. This structure (simplified in Fig. 1) provides many insights. One challenging conclusion relates to the positioning of the Cicereae as an outgroup to both the Viceae and Trifoleae with paraphyly in the latter. One of the few intra-generic divergence times given by Lavin *et al.* (2005), within the genus *Cicer*, has a basal divergence at 14.8 My. This is greater than several inter-generic divergence times; some genera diverged much more recently, and presumably species more recently again. Jing *et al.* (2005) obtained estimates of the age of alleles segregating within *Pisum* and obtained values between 1.5 and 5 My. Together, these pieces of information show how taxonomic divisions can represent rather different divergence times within and between species or genera depending on context.

The insights from molecular taxonomy allow us to form evolutionary hypotheses more clearly, and underscore the coherence and unity of legumes as a group. This coherence is reflected in our understanding of the conservative nature of the structure of legume nuclear genomes (Choi *et al.*, 2004; Kaló *et al.*, 2004; George *et al.*, 2008; Hougaard *et al.*, 2008; Timko *et al.*, 2008; Bertioli *et al.*, 2009; Muchero *et al.*, 2009) and conservation of gene function (Domoney *et al.*, 2006). For example, the *Dt1* gene of soybean (Liu *et al.*, 2010) discussed by Nelson in this volume corresponds to *Det* in pea (Foucher *et al.*, 2003) and most likely the equivalent growth habit determinant in faba bean (Avila *et al.*, 2006).

The ease with which molecular markers can be developed for crop species has enabled us to obtain high resolution descriptions of the pattern of genetic variation within species. In general, we also find that genetic variation is unevenly partitioned with respect to eco-geographical factors, so an emerging concern is of the extent to which the consequent correlations reflect local adaptation and hence may help direct us to specific

* Corresponding author. E-mail: noel.ellis@bbsrc.ac.uk
†Present address: Institute of Biological, Environmental & Rural Sciences, Aberystwyth University, Gogerddan, Aberystwyth, Ceredigion, SY23 3EB. E-mail: noe2@aber.ac.uk

https://doi.org/10.1017/S1479262110000432 Published online by Cambridge University Press

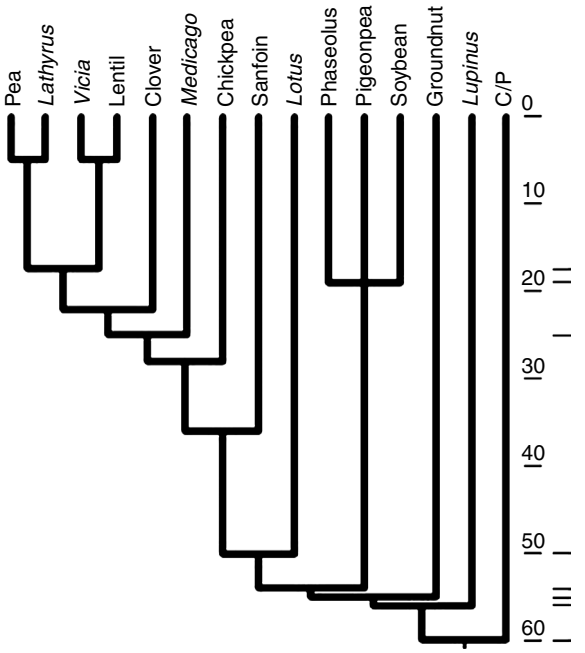


Fig. 1. Simplified legume taxonomy from Lavin *et al.* (2005); a time-line is indicated to the right, and divergences dated by Lavin *et al.* (2005) are indicated to the right. The divergence time for chickpea is between 24.7 and 32.1 My and for Sanfoin is between 32.1 and 39 My based on the average of flanking nodes. The position of *Cajanus* (pigeonpea) is not clear from Lavin *et al.,* but it groups with the Phaseoleae that include Phaseolus and *Glycine* according to Doyle and Lucknow (2003).

subsets of germplasm for different purposes. It seems that the genotypic description of germplasm collections is a powerful tool, especially for inbreeding species. The extent to which genotype data can remain associated with germplasm collections is addressed by Nelson's article in this volume, and has caused much concern widely. If accessions are heterogeneous, then what is the detection limit for particular alleles and how does this propagate to 'missing data' in subsequent analyses? One approach has been to replicate germplasm collections from single individuals which were the material for genotypic analysis (Jing *et al.*, 2010), but this may not be appropriate for much larger collections.

The concept of 'core' and 'composite' collections, for example in the discussion of the pigeonpea collection by Upadhyaya *et al.*, is another recurring theme. This type of sampling represents a way of simplifying analyses, but there is a diversity of opinions on how best to accomplish this simplification. Simplified sets of germplasm have the advantage that they can be managed more easily and phenotyping experiments become feasible, but they have the disadvantage that some variation is lost. It seems likely that a combination of information on genotype and provenance may in the future aid in the definition of different sets of material appropriate for different purposes.

This introduction has discussed the linked themes of phylogenetic relationships between, and diversity within, crop species, which represents a combination central to much of N. I. Vavilov's work. His interest in legumes is presented in several of the articles in this volume, and indeed the genus that bears his name is also discussed by Smykal *et al.* It is sobering to think of the insights he brought to biology, and especially to plant genetics, with what seems relatively little data. What would he have made of our data and concerns? It seems to me that this creates two distinct challenges. One is to see out of the fog of data to try to find clear and general principles. The other is to use the precision to hand for practical purpose in crop improvement.

References

Avila CN, Nadal S, Moreno MT and Torres AM (2006) Development of a simple PCR-based marker for the determination of growth habit in *Vicia faba* L. using a candidate gene approach. *Molecular Breeding* 17: 185–190.

Bertioli DJ, Moretzsohn MC, Madsen LH, Sandal N, Leal-Bertioli SCM, Guimarães PM, Hougaard BK, Fredslund J, Schauser L, Nielsen AM, Sato S, Tabata S, Cannon SB and Stougaard J (2009) An analysis of synteny of *Arachis* with *Lotus* and *Medicago* sheds new light on the structure, stability and evolution of legume genomes. *BMC Genomics* 10: 45.

Chase MW, Soltis DE, Olmstead RG, Morgan D, Les DH, Mishler BD, Duvall MR, Price RA, Hills HG, Qiu Y-L, Kron KA, Rettig JH, Conti E, Palmer JD, Manhart JR, Sytsma KJ, Michaels HJ, Kress WK, Karol KG, Clark WD, Hedren M, Gaut BS, Jansen RK, Kim K-J, Wimpee CF, Smith JF, Furnier GR, Strauss SH, Xiang Q-Y, Plunkett GM, Soltis PS, Swensen SM, Williams SE, Gadek PA, Quinn CJ, Eguiarte LE, Golenberg E, Learn GH, Graham SW, Barrett SCH, Dayanandan S and Albert VA (1993) Phylogenetics of seed plants: an analysis of nucleotide sequences from the plastid gene *rbcL*. *Annals of the Missouri Botanical Garden* 80: 528–580.

Choi H-K, Mun J-H, Kim D-J, Zhu H, Baek J-M, Mudge J, Roe B, Ellis N, Doyle J, Kiss GB, Young ND and Cook D (2004) Estimating genome conservation between crop and model legume species. *PNAS* 101: 15289–15294.

Domoney C, Duc G, Ellis THN, Ferrándiz C, Firnhaber C, Gallardo K, Hofer J, Kopka J, Küster H, Madueño F, Munier-Jolain NG, Mayer K, Thompson R, Udvardi M and Salon C (2006) Genetic and genomic analysis of legume flowers and seeds. *Current Opinion in Plant Biology* 9: 133–141.

Doyle JJ and Luckow MA (2003) The Rest of the Iceberg. Legume Diversity and Evolution in a Phylogenetic Context. *Plant Physiology* 131: 900–910.

Foucher F, Morin J, Courtiade J, Cadioux S, Ellis N, Banfield MJ and Rameau C (2003) Determinate and late flowering are two terminal flower 1/centroradialis homologues controlling two distinct phases of flowering initiation and development in pea. *Plant Cell* 15: 2742–2754.

https://doi.org/10.1017/S1479262110000432 Published online by Cambridge University Press

George J, Sawbridge TI, Cogan NOI, Gendall AR, Smith KF, Spangenberg GC and Forster JW (2008) Comparison of genome structure between white clover and *Medicago truncatula* supports homoeologous group nomenclature based on conserved synteny. *Genome* 51: 905–911.

Hougaard BK, Madsen LH, Sandal N, Moretzsohn MdC, Fredslund J, Schauser L, Nielsen AM, Rohde T, Sato S, Tabata S, Bertioli J and Stougaard J (2008) Legume anchor markers link syntenic regions between *Phaseolus vulgaris*, *Lotus japonicus*, *Medicago truncatula* and *Arachis*. *Genetics* 179: 2299–2312.

Jing R, Knox MR, Lee JM, Vershinin AV, Ambrose M, Ellis THN and Flavell AJ (2005) Insertional polymorphism and antiquity of *PDR1* retrotransposon insertions in *Pisum* species. *Genetics* 171: 741–752.

Kaló P, Seres A, Taylor SA, Jakab J, Kevei Z, Kereszt A, Endre G, Ellis THN and Kiss GB (2004) Comparative mapping between *Medicago sativa* and *Pisum sativum*. *Molecular Genetics and Genomics* 272: 235–246.

Lavin M, Herendeen PS and Wojciechowski MF (2005) Evolutionary rates analysis of leguminosae implicates a rapid diversification of lineages during the tertiary. *Systematic Biology* 54: 575–594.

Liu B, Watanabe S, Uchiyama T, Kong F, Kanazawa A, Xia Z, Nagamatsu A, Arai M, Yamada T, Kitamura K, Masuta C, Harada K and Abe J (2010) The soybean stem growth habit gene *Dt1* is an ortholog of Arabidopsis terminal flower 1. *Plant Physiology* 153: 198–210.

Muchero W, Diop NN, Bhat PR, Fenton RD, Wanamaker S, Marti Pottorff M, Hearne S, Cisse N, Fatokun C, Ehlers JD, Roberts PA and Close TJ (2009) A consensus genetic map of cowpea [*Vigna unguiculata* (L) Walp.] and synteny based on EST-derived SNPs. *PNAS* 106: 18159–18164.

Timko MP, Rushton PJ, Laudeman TW, Bokowiec MT, Chipumuro E, Cheung F, Town CD and Chen X (2008) Sequencing and analysis of the gene-rich space of cowpea. *BMC Genomics* 9: 103.

https://doi.org/10.1017/S1479262110000432 Published online by Cambridge University Press