Naturally occurring diversity helps to reveal genes of adaptive importance in legumes

Laurent Gentzbittel1,2*, Stig U. Andersen3, Cécile Ben1,2, Martina Rickauer1,2, Jens Stougaard3 and Nevin D. Young4,5

1 EcoLab Laboratoire Écologie Fonctionnelle et Environnement, Institut National Polytechnique de Toulouse, Ecole Nationale Supérieure Agronomique de Toulouse, Université Fédérale de Toulouse, Castanet Tolosan, France, 2 EcoLab Laboratoire Écologie Fonctionnelle et Environnement, Centre National de la Recherche Scientifique, Castanet Tolosan, France, 3 Department of Molecular Biology and Genetics, Centre for Carbohydrate Recognition and Signalling, Aarhus University, Aarhus, Denmark, 4 Department of Plant Pathology, University of Minnesota, St. Paul, MN, USA, 5 Department of Plant Biology, University of Minnesota, St. Paul, MN, USA

Environmental changes challenge plants and drive adaptation to new conditions, suggesting that natural biodiversity may be a source of adaptive alleles acting through phenotypic plasticity and/or micro-evolution. Crosses between accessions differing for a given trait have been the most common way to disentangle genetic and environmental components. Interestingly, such man-made crosses may combine alleles that never meet in nature. Another way to discover adaptive alleles, inspired by evolution, is to survey large ecotype collections and to use association genetics to identify loci of interest. Both of these two genetic approaches are based on the use of biodiversity and may eventually help us in identifying the genes that plants use to respond to challenges such as short-term stresses or those due to global climate change. In legumes, two wild species, *Medicago truncatula* and *Lotus japonicus*, plus the cultivated soybean (*Glycine max*) have been adopted as models for genomic studies. In this review, we will discuss the resources, limitations and future plans for a systematic use of biodiversity resources in model legumes to pinpoint genes of adaptive importance in legumes, and their application in breeding.

Keywords: *Medicago truncatula*, *Lotus japonicus*, *Glycine max*, GWAS, genetics, whole genome sequencing, gene expression

Introduction

The legume family (*Leguminosae*) is second only to cereals in economic and nutritional value. Grain legumes, such as common bean (*Phaseolus vulgaris*), lentil (*Lens culinaris*), or chickpea (*Cicer arietinum*) provide on average 33% of human’s dietary nitrogen and up to 60% in developing countries (O’Rourke et al., 2014). Altogether, grain legumes are grown on 115 million ha, with soybean for feed and food grown on 111 million ha, representing 276 million tons. The major fodder legume, alfalfa, is cultivated over 15 million ha for 340 million tons of forage.

The legume family consists of approximately 20,000 species (Doyle and Lucknow, 2003). Legumes are often pioneer plants improving soil fertility and moderating harsh conditions.
as demonstrated with *Lotus corniculatus* (Esperschütz et al., 2011). **Figure 1** summarizes the centers of origin for several important cultivated and model legumes. Legumes establish symbiotic interactions with both rhizobia (the *Rhizobium*-legume symbiosis, RL) and arbuscular mycorrhizal (AM) fungi leading to formation of nitrogen-fixing nodules and phosphate acquiring mycorrhiza (Oldroyd, 2013). These traits play a vital role in ecosystems and sustainable crop production, and are central for efforts to decrease dependence on commercial fertilizers and substitute imported protein feeds with locally produced legume crops (Voisin et al., 2014). Aiming to uncover the genetic basis for these features, *Medicago truncatula* and *Lotus japonicus*, have come to supplement *Arabidopsis*, maize and rice as plant models. Through a mutant approach a core set of RL symbiosis genes has been identified in the two model legumes (Kouchi et al., 2010; Gough and Cullimore, 2011). Virtually all current knowledge about the perception of rhizobial signals, transduction and organogenesis originates from these studies. The RL and AM signaling pathways overlap and the genetic basis of this common symbiosis pathway comprises more than 10 genes (Parniske, 2008). Several mutants affected in RL symbiosis also show altered responses to root pathogens (Rey et al., 2013; Ben et al., 2013b) suggesting crosstalk between signaling pathways in symbiosis and diseases.

Translational biology (from models to crops) is still in its infancy. This could be due to lack of genetic variability, where the crop’s orthologous gene is not polymorphic and thus not amenable to breeding. Strategies to introduce heterologous genes into crops may encounter technical difficulties, or societal issues such as consumers’ acceptance. Genes identified in mutant-based approaches may also be unrelated to adaptation and responses to biotic and abiotic stresses encountered in agriculture or natural environments. Initial phenotypic screens may not have been set up to target “adaptive” traits such as seed yield, biomass, or fitness. Nonetheless, legume crops do benefit from models: the RCT1 gene from *M. truncatula* provides resistance to anthracnose in alfalfa (Yang et al., 2008). Wild species also contributed to the improvement of crops: a high-protein allele has been introgressed into soybean ([*Glycine max* (L.) Merr.] from its wild progenitor, *Glycine soja* Sieb. and Zucc (Sebolt et al., 2000; Vaughn et al., 2014). Many legume crops have limited genetic diversity due to bottlenecks from domestication and selective breeding (Gross and Olsen, 2010; Roorkiwal et al., 2014). Proper characterization and evaluation of germplasm collections as sources of adaptive alleles, and their utilization in breeding are often limited or neglected (Smykal et al., 2015). However, the anticipated climate change may require explicit efforts to breed for local adaptation. The use of wild relatives of cultivated legumes crops has received limited attention as well. As a consequence, selection of new cultivars and improving production technologies for food legume crops are not proceeding at the same pace as for cereal crops.

*M. truncatula* and *L. japonicus* as members of the *Hologalegina* clade (which contains cultivated legumes such as pea, chickpea), and *G. max* as a representative of the *Phaseoloid* clade (containing mungbean and soybean) can serve as models for cool season and tropical season legume crops, respectively (Doyle and Lucknow, 2003; Zhu et al., 2005). The high degree of synteny between these model plants and their legume crops relatives can be exploited to improve genetic maps and identify candidate genes for agronomic traits in less well-characterized crops, as has been described for leaf size in cowpea (Potorff et al., 2012), drought adaptation traits in faba bean (Kharzaei et al., 2014) or symbiotic genes in pea (Novak et al., 2012). In this Perspective, we will discuss the resources, limitations and future plans for a systematic use of biodiversity resources in model legumes to pinpoint genes of adaptive importance in legumes.

**FIGURE 1** | Putative centers of origin of major legume crops and model legumes.
Exploring the Natural Diversity of Medicago truncatula

Medicago truncatula is closely related to alfalfa. Together with L. japonicus, M. truncatula has been essential for studies of RL and AM symbioses. In addition, agronomic and quality traits (Julier et al., 2007), including drought and salinity tolerance (Friesen et al., 2014), seed development and composition (Vandecasteele et al., 2011), and disease resistance (e.g., Ben et al., 2013) have been targeted. Consequently, this species has an excellent array of germplasm, molecular, and genomic resources.

Studies of Medicago populations focused on collections from the Mediterranean basin, its center of diversity. Work in France led to the establishment of extensive germplasm resources (Thoquet et al., 2002), including accesses covering M. truncatula's natural range plus strategically selected biparental crosses and derived recombinant inbred lines. Additional material is available from SARDI (Ellwood et al., 2006) and USDA-GRIN. Another valuable germplasm resource is the TntI insertion population created at the Noble Foundation, which provides knockout mutants for reverse genetics and characterization of candidate genes (Pisarri et al., 2012). The genome sequence released in 2011 (Young et al., 2011) and refined in 2014 (Tang et al., 2014) established M. truncatula as an outstanding system for genome studies. Altogether the Mt4.0 assembly covers 360 Mb and gene annotation uncovers 31,661 high confidence predicted genes.

The underlying genome architecture was first described in Branca et al. (2011) using 26 accessions broadly spanning M. truncatula's range, at 30x sequence coverage. This uncovered four to six SNPs/kb and LD decays reaching half its initial value within 3–4 kb, similar to that of Arabidopsis (Kim et al., 2007).

More than 20,000 annotated genes from 56 accessions were used to identify targets of positive selection (Paape et al., 2013). Around 1% of sampled genes harbored a signature of positive selection, while 50–75% of non-synonymous polymorphisms were subject to purifying selection. Among putative targets of selection were genes involved in defense against pathogens and herbivores and in relationship with rhizobial symbionts.

Local adaptation and adaptive clines were examined in 202 accessions at 2 million SNPs, identifying loci responsible for adaptation to climatic gradients: annual mean temperature (AMT), precipitation in the wettest month (PWM), and isothermality (Yoder et al., 2014). The strongest associations tagged genome regions containing genes with predicted roles in tolerance to temperature, drought, herbivores or pathogens. The candidate loci were further tested and validated using climate-controlled tests. For AMT and PWM, a history of soft selective sweeps acting on loci underlying adaptation was indicated.

GWAS based on 6 million SNPs, identified by re-sequencing 226 M. truncatula accessions, revealed candidate genes underlying phenotypic variation in several plant functional traits (Stanton-Geddes et al., 2013). For flowering time and trichome density, peaks were associated with well-supported candidates (MtFD for flowering time; unshaven in the case of trichome density). For rhizobium symbiosis, previously characterized nodulation genes (SERK2, MtnodGRP3, MtMMPL1, NFP, CaML3, MtnodGRP3A) were confirmed, and novel loci were identified with annotation and/or expression profiles that supported a role in nodule formation.

Population genomics revealed candidate regions associated with local adaptation (Friesen et al., 2010, 2014) to saline environments, by searching for SNP that assorted by population i.e., two populations from saline environments vs. two from non-saline environments. The analyses pinpointed signaling pathways for biotic stress tolerance involving ABA and MeJA, production of putative osmoprotectant, and candidates linked to biotic interactions. A strategy for salt stress avoidance through early flowering was suggested, specifically a non-synonymous SNP that changes a highly conserved amino acid in the Medicago ortholog of CONSTANS (Pierre et al., 2011).

M. truncatula is a host for leaf and soil-borne pathogens (Ameline-Torregrosa et al., 2007; Ramírez-Suero et al., 2010; Risplid and Rubiales, 2014). Quantitative resistance loci for root diseases have been described for the interaction with Ralstonia solanacearum (Ben et al., 2013a), the oomycete Aphanomyces euteiches (Ae) (Djébali et al., 2009, Pilet-Nayel et al., 2009), and Verticillium alfafae (Ben et al., 2013b). Bonhomme et al. (2014) utilized GWAS to map loci associated with Ae resistance. This identified several candidate genes and pinpointed two major loci on chromosome 3, that co-located with previously described QTLs (Djébali et al., 2009; Pilet-Nayel et al., 2009). Intriguingly, the resistance allele of the candidate gene seemed to be the ancestral form, although the pathogen and the resistant host plant do not occur naturally in the same geographical regions. Another study of Ae resistance with 136 lines from 14 Tunisian populations, not naturally exposed to the oomycete, suggested that resistance can occur as a by-product of adaptation to water stress (Djébali et al., 2013).

Clearly, the evaluation of naturally occurring variations in M. truncatula, making use of its well-characterized biodiversity and feature-rich genomic tools, is becoming a powerful strategy of investigation. Identification of candidate genes for major traits is likely to expand rapidly.

Lotus Resources for Functional Analysis of Natural Diversity

Several Lotus species are cultivated as forage legumes, especially in grassland areas characterized by harsh environments, including windy and saline coastal climates (Escaray et al., 2012). The genetic background for such plasticity and adaptation to adverse conditions in poor soils, flooding, drought, salt and other abiotic stresses is important for plant production in a world facing climate changes. Model plants are essential for unequivocal identification of genetic regulators, and one of the diploid and self-fertile Lotus species, L. japonicus, was proposed in 1992 as a model system for classical and molecular genetics (Handberg and Stougaard, 1992). Since then, genetic and physical maps, F2 and recombinant inbred line populations, transformation protocols and a reference genome sequence have been established (Hayashi et al., 2001; Kawaguchi et al., 2001, 2005; Lohar et al., 2001; Sandal et al., 2002, 2012; Sato et al., 2008). These resources
have facilitated molecular cloning of central genetic components required for endosymbiosis with rhizobia and arbuscular mycorrhiza (Schauser et al., 1999; Stracke et al., 2002; Radutoiu et al., 2003; Imaizumi-Anraku et al., 2005; Yano et al., 2008), positioning L. japonicus as a major model system for legume research. In a broader context, the L. japonicus resources were also exploited for comparative genomic approaches in a number of crop legumes, including pea, bean, and lupin, and for gene cloning in pea (Choi et al., 2004; Stracke et al., 2004; Fredslund et al., 2005; Hougard et al., 2008; Li et al., 2010; McConnell et al., 2010; Nelson et al., 2010; Humphry et al., 2011; Krussel et al., 2011; Cruz-Izquierdo et al., 2012; Shirasawa et al., 2014).

Genetic analysis of natural variation within L. japonicus has now also been initiated. First, biparental populations have been used for QTL mapping of agronomic traits, host specificity, and nitrogen fixation efficiency using the three central experimental genotypes MG-20, Gifu and L. bertii (Gondo et al., 2007; Sandal et al., 2012; Tominaga et al., 2012), and diversity information has been exploited for evaluation the impact of Nod factor perception on rhizobium host range (Radutoiu et al., 2007; Bek et al., 2010). Second, a collection of ~200 L. japonicus accessions is managed and distributed by the Japanese National Bioresource Program (Hashiguchi et al., 2012). Phenotypic characterization has been carried out for some of these accessions (Kai et al., 2010) and so far 130 of them have been re-sequenced to facilitate GWAS analysis (Sato and Andersen, unpublished). There is a strong interest in the community for taking advantage of these new resources, and the accessions have been phenotyped for flowering time, trichome phenotype, nodulation capacity, salt tolerance, nematode resistance and root growth traits (Poch et al., 2007; Kubo et al., 2009; Gossmann et al., 2012; Wakabayashi et al., 2014; and unpublished results).

To take full advantage of GWAS-based genetic analysis and bring it to fruition in terms of functional characterization of genes, additional resources are needed. Expression data, high quality annotation and homology information is available for L. japonicus, facilitating quick selection of the most promising candidate genes for further investigation (Gonzales et al., 2005; Hogslund et al., 2009; Li et al., 2012; Verdier et al., 2013). Likewise, easy access to insertion mutants is critical to validate candidate gene involvement in a phenotype. Here, L. japonicus offers access to a unique resource in the non-transgenic LORE1 collection, which currently holds 40,000 annotated lines, with an additional 60,000 lines already characterized and soon to be released, bringing the total of annotated insertions to more than 500,000 (Urbanski et al., 2012; Andersen, unpublished). Together with the TILLING population this offers a comprehensive platform for mutant analysis (Perry et al., 2003).

The sum of these resources makes L. japonicus attractive as a model species, and the availability of non-transgenic insertion mutants offers the unique possibility of phenotyping L. japonicus and other legume diversity panels under similar conditions to estimate the degree of conservation of the genetic networks governing phenotypic responses across legumes and facilitating functional follow-up by characterization of candidate LORE1 mutant phenotypes under field conditions.

### Soybean Benefits from Comparison with its Wild Progenitors

Soybean is the most important legume crop and one of the largest global sources of vegetable oil and protein for people and livestock (Graham and Vance, 2003). Not surprisingly, the genome resource for soybean first published in 2010 is among the best for plants (Schmutz et al., 2010). For the initial assembly, more than 950 Mbp of the overall 1115 Mbp genome were completed through 8X Sanger shotgun sequencing and 98% could be anchored to specific chromosomal positions with few gaps. With this outstanding resource as a starting point, genome re-sequencing, genome-wide scans for selection, and GWAS have been fruitful in Glycine sp.

As a community-wide resource for genotyping a “SoySNP50K iSelect BeadChip” (Song et al., 2013) was designed from six cultivated and two wild soybean, aligned to the soybean genome reference. A total of 47,337 SNP calls were identified when tested against 96 landraces, 96 elite cultivars, and 96 wild accessions. This chip is used in many recent soybean GWAS publications.

Protein and oil content in the soybean seed are of special interest. In a GWAS analysis of 298 genotypes, Hwang et al. (2014) uncovered 40 associated SNPs in 17 different regions. One region was a long-sought protein/oil QTL on linkage group 1 (chromosome 20). This gene has been the target of QTL mapping research for many years, but earlier (biparental) studies struggled to localize the target gene to a region smaller than 8 Mbp (Bolon et al., 2011). In a second protein/oil GWAS study, Vaughn et al. (2014) discovered a more diagnostic and presumably more tightly linked locus as well as additional novel seed composition loci.

Further GWAS studies focused on abiotic and biotic stress. Concerning water deficit tolerance, Dhanapal et al. (2015) characterized carbon isotope ratios for 375 genotypes across four environments and 2 years, and described SNP tagging of 21 loci. Tolerance to low phosphorous is another important trait and GWAS enabled Zhang et al. (2014) to identify six regions associated with P efficiency. Closer examination of a region on chromosome 8 across 192 soybean accessions revealed a candidate, GmACP1 (an acid phosphatase). Favorable alleles and haplotypes of GmACP1 are associated with higher enzyme activity. Finally, resistance to the fungal pathogen Fusarium virguliforme, the cause of sudden death syndrome (SDS), was analyzed through GWAS in two association panels of elite cultivars by Wen et al. (2014). A total of 20 loci underlying SDS resistance were discovered, seven in regions previously described and 13 novel ones. One of these loci overlapped a previously cloned SDS resistance gene, Rfs2.

The availability of genome resources has enabled whole genome re-sequencing, focusing on soybean domestication from wild soybean (G. soja). For example, G. soja has been found to differ from the reference genome by only 0.31% (Kim et al., 2010). Re-sequencing 17 wild and 14 cultivated soybean genomes revealed unusually high levels of linkage disequilibrium (LD) with LD blocks greater than 1 Mb in cultivated soybean (Lam et al., 2010). Based on Fst scans across the genome, putative domestication regions (areas of high Fst) could be identified.
Detailed mapping pinpointed smaller regions associated with known domestication traits such as twinning and stem elongation. By re-sequencing 55 accessions, Li et al. (2013) show that selection during early domestication led to more pronounced reduction in genetic diversity than the move from landraces to elite cultivars. Clusters of selection hotspots were observed involving 4.38% of total annotated genes. Finally, Li et al. (2014) initiated the first step toward a soybean pan-genome, generating de novo genome sequences independent of the soybean (G. max Williams–82) reference. This is important for characterizing structural variation, including copy number variation and for identifying core vs. dispensable portions of the soybean genome.

**Making the Most of Legumes – Collaborations across Species**

To date, research efforts have focused on developing the resources, reference genome sequences and catalogs of natural diversity and germplasm, required to carry out association studies within a species. With the successful establishment of these prerequisites, the time has come to consider the possibilities offered by the resources, along with those emerging in multiple legume species. While GWAS is a powerful tool for genetic analysis and causal allele identification, it also comes with inherent challenges. Inaccuracies in geno- and phenotyping combined with population sampling and structure can all lead to false positive signals. However, these confounding factors are species-specific, as the legume resources have largely been developed independently for each species. As such, genotype-phenotype associations detected recurrently in at least two legume species would therefore be strongly mutually supportive.

Leveraging these potential gains in GWAS candidate gene confidence will require a new paradigm for GWAS phenotyping, where multi-site, multi-year phenotyping strategies are combined with multi-species trials. Approaches where multiple legume species are grown side-by-side for phenotyping at multiple geographical locations will not be relevant for all traits, but obvious traits suitable for analysis are abiotic and broad host-range biotic stresses where similar adaptive strategies could be in play across species. Such phenotyping strategies have the potential to break new ground in the understanding of environmental adaptation, but organization of international consortia that could successfully implement these approaches would require a high level of (re-)organization, for researchers, breeders, and funding bodies.

**Acknowledgments**

We apologize to all authors whose relevant work was not cited due to space constraints.

**References**

Ameline-Torregrosa, C., Cazaux, M., Danesh, D., Chardon, F., Cannon, S. B., Esquerre-Tugayé, M.-T., et al. (2007). Genetic dissection of resistance to anthracnose and powdery mildew in Medicago truncatula. Mol. Plant Microbe Interact. 21, 61–69. doi: 10.1094/MPMI-21-1-0061

Bek, A. S., Sauer, J., Thygesen, M. R., Dvuss, J. Ő., Petersen, B. O., Thirup, S., et al. (2010). Improved characterization of nod factors and genetically based variation in LysM receptor domains identify amino acids expendable for nod factor recognition in Lotus spp. Mol. Plant Microbe Interact. 23, 58–66. doi: 10.1094/MPMI-23-1-0058

Ben, C., Debellé, F., Berges, H., Bellec, A., Jardinard, M.-F., Anson, P., et al. (2013a). MtQRRS1, an R-locus required for Medicago truncatula quantitative resistance to Ralstonia solanacearum. New Plant. 199, 758–772. doi: 10.1111/nph.12299

Ben, C., Toueni, M., Montanari, S., Tardin, M.-C., Ferval, M., Negahi, A., et al. (2013b). Natural diversity in the model legume Medicago truncatula allows identifying distinct genetic mechanisms conferring partial resistance to Verticillium wilt. J. Exp. Bot. 64, 317–332. doi: 10.1093/jxb/ers337

Bolon, Y.-T., Haun, W. J., Xu, W. W., Grant, D., Stacey, M. G., Nelson, R. T., et al. (2011). Phenotypic and genomic analyses of a fast neutron mutant population resource in soybean. Plant Physiol. 156, 240–253. doi: 10.1104/pp.110.170811

Bonhomme, M., André, O., Badis, Y., Ronfort, J., Burgarella, C., Chardon, F., Jaunneau, N., et al. (2014). High-density genome-wide association mapping implicates an F-box encoding gene in Medicago truncatula resistance to Aphanomyces euteiches. New Plant. 201, 1328–1342. doi: 10.1111/nph.12611

Branca, A., Paape, T. D., Zhou, P., Brisikine, R., Farmer, A. D., Mudge, J., et al. (2011). Whole-genome nucleotide diversity, recombination, and linkage disequilibrium in the model legume Medicago truncatula. Proc. Natl. Acad. Sci. U.S.A. 108, E864–E870. doi: 10.1073/pnas.110432108

Choi, H.-K., Mun, J.-H., Kim, D.-J., Zhu, H., Baek, J.-M., Mudge, J., et al. (2004). Estimating genome conservation between crop and model legume species. Proc. Natl. Acad. Sci. U.S.A. 101, 15289–15294. doi: 10.1073/pnas.0402251101

Cruz-Izquierdo, S., Avila, C. M., Satovic, Z., Palomino, C., Gutierrez, N., Ellwood, S. R., et al. (2012). Comparative genomics to bridge Vicia faba with model and closely-related legume species: stability of QTLs for flowering and yield-related traits. Theor. Appl. Genet. 125, 1767–1782. doi: 10.1007/s00122-012-2152-1

Dhanapal, A. P., Ray, J. D., Singh, S. K., Hoyos-Villegas, V., Smith, J. R., Purcell, L. C., et al. (2015). Genome-wide association study (GWAS) of carbon isotope ratio (δ13C) in diverse soybean [Glycine max (L.) Merr.] genotypes. Theor. Appl. Genet. 128, 73–91. doi: 10.1007/s00122-014-2413-9

Djebali, N., Aribi, S., Taamalli, W., Arraooudi, S., Aouani, M. E., and Badri, M. (2013). Natural variation of Medicago truncatula resistance to Aphanomyces euteiches. Eur. J. Plant Pathol. 135, 831–843. doi: 10.1007/s10658-012-0127-x

Djebali, N., Jaunneau, A., Ameline-Torregrosa, C., Chardon, F., Jaunneau, V., Mathé, C., et al. (2009). Partial resistance of Medicago truncatula to Aphanomyces euteiches is associated with protection of the root stele and is controlled by a major QTL rich in proteasome-related genes. Mol. Plant Microbe Interact. 22, 1043–1055. doi: 10.1094/MPMI-22-9-1043

Doyle, J. J., and Lucknow, M. A. (2003). The rest of the iceberg. Legume diversity and evolution in a phylogenetic context. Plant Physiol. 131, 900–910. doi: 10.1104/pp.102.018150

Ellwood, S. R., D’Souza, N. K., Kamphuis, L. G., Burgess, T. L., Nair, R. M., and Oliver, R. P. (2006). SSR analysis of the Medicago truncatula SARCO core collection reveals substantial diversity and unusual genotype dispersal throughout the Mediterranean basin. Theor. Appl. Genet. 112, 977–983. doi: 10.1007/s00122-005-0202-1

Escary, F. J., Menendez, A. B., Gárriz, A., Pieckenstain, F. L., Estrella, M. J., Castagno, L. N., et al. (2012). Ecological and agronomic importance of the plant genus Lotus. Its application in grassland sustainability and the amelioration of constrained and contaminated soils. Plant Sci. Int. J. Exp. Plant Biol. 182, 121–133. doi: 10.1016/j.plantsci.2011.03.016

Epperschütz, J., Weid, G., Schreiner, K., Buegger, F., Munch, J. C., and Schloter, M. (2011). Incorporation of carbon from decomposing litter of two pioneer plant species into microbial communities of the detritusphere: C fluxes from litter into microorganisms in initial ecosystems. FEMS Microbiol. Lett. 320, 48–55. doi: 10.1111/j.1574-6968.2011.02286.x

Fredslund, J., Schauer, L., Madsen, L. H., Sandal, N., and Stougaard, J. (2005). PiriFi: using a multiple alignment of related sequences to find primers for

---

**Biodiversity for gene identification in legumes**

Frontiers in Plant Science | www.frontiersin.org 5 April 2015 Volume 6 Article 269
Parniske, M. (2008). *Arbuscular mycorrhiza*: the mother of plant root endosymbioses. *Nat. Rev. Microbiol.* 6, 763–775. doi:10.1038/nrmicro1987

Perry, J. A., Wang, T. L., Welham, T. J., Gardner, S., Pike, J. M., Yoshiida, S., et al. (2003). A TILLING reverse genetics tool and a web-accessible collection of mutants of the legume *Lotus japonicus*. *Plant Physiology* 131, 866–871. doi:10.1104/pp.102.027345

Pierre, J.-P., Bogard, M., Herrmann, D., Huyghe, C., and Julier, B. (2011). A CONSTANT-like gene candidate that could explain mosy of the genetic variation for flowering date in *Medicago truncatula*. *Mol. Breeding* 28, 25–35. doi:10.1007/s11032-010-9457-6

Pilet-Nayel, M.-L., Prospéri, J.-M., Hamon, C., Lesné, A., Lecointe, R., Le Goff, I., et al. (2009). AER1, a major gene conferring resistance to *Aphanomyces euteiches* in *Medicago truncatula*. *Phytopathology* 99, 203–208. doi:10.1094/PHYTO-99-2-0203

Pislariu, C. I., Murray, J. D., Wen, J., Cosson, V., Muni, R. R. D., Wang, M., et al. (2012). *A Medicago truncatula* tobacco retrotransposon insertion mutant collection with defects in nodule development and symbiotic nitrogen fixation. *Plant Physiology* 159, 1686–1699. doi:10.1104/pp.112.197061

Poch, H. L. C., López, R. H. M., and Clark, S. J. (2007). Ecotypes of the model legume *Lotus japonicus* vary in their interaction phenotypes with the root-knot nematode *Meloidogyne incognita*. *Ann. Bot.* 99, 1223–1229. doi:10.1093/aob/mcm058

Potorff, M., Ehlers, J. D., Fatoku, C., Roberts, P. A., and Close, T. J. (2012). Leaf morphology incoaprewa (*Vigna unguiculata* L.): QTL analysis, physical mapping and identifying a candidate gene using synteny with model legume species. * BMC Genomics* 13:234. doi:10.1186/1471-2164-13-234

Radoutioiu, S., Madsen, L. H., Madsen, E. B., Felle, H. H., Umehara, Y., Gronlund, M., et al. (2003). Plant recognition of symbiotic bacteria requires two LysM receptor-like kinases. *Nature* 425, 585–592. doi:10.1038/nature02039

Ramiezu, S., Madsen, L. H., Madsen, E. B., Felle, H. H., Umehara, Y., Gronlund, M., et al. (2003). Plant recognition of symbiotic bacteria requires two LysM receptor-like kinases. *Nature* 425, 585–592. doi:10.1038/nature02039

Rey, T., Nars, A., Bonhomme, M., Bottin, A., Huguet, S., Balzerque, S., et al. (2013). NFP, a LysM protein controlling Nod factor perception, also intervenes in *Medicago truncatula* resistance to pathogens. *New Phytol.* 198, 875–886. doi:10.1111/nph.12198

Rispens, N., and Rubiales, D. (2014). Identification of sources of quantitative resistance to *Fusarium oxysporum* f. sp. medicaginis in *Medicago truncatula*. *Plant Pathology* 63, 986–973. doi:10.1111/ppa.12212

Roorkiwal, M., von Wettberg, E. J., Upadhyaya, H. D., Warschefsky, E., Rathore, A., Krishnakumar, M., and Varshney, R. K. (2014). Exploring germplasm diversity to understand the domestication process in *Lotus*. *Front. Plant Sci.* 5:109. doi:10.3389/fpls.2014.00189-y

Song, Q., Hyten, D. L., Jia, G., Quigley, C. V., Fickus, E. W., Nelson, R. L., et al. (2013). Development and evaluation of *SoySNP50K*, a high-density genotyping array for soybean. *PloS ONE* 8:e54985. doi:10.1371/journal.pone.0054985

Stanton-Geddes, J., Pape, T., Epstein, B., Briskine, R., Yoder, J., Mudge, J., et al. (2013). Candidate Genes and genetic architecture of symbiotic and agromonic traits revealed by whole-genome, sequence-based association genomics in *Medicago truncatula*. *PloS ONE* 8:e65688. doi:10.1371/journal.pone.0065688

Stracke, S., Knister, C., Yoshida, S., Mulder, L., Sato, S., Kaneko, T., et al. (2002). A plant receptor-like kinase involved in both bacterial and fungal symbiosis. *Nature* 417, 959–962. doi:10.1038/nature00841

Stracke, S., Sato, S., Sandal, N., Koyama, M., Kaneko, T., Tabata, S., et al. (2004). Exploitation of colinear relationships between the genomes of *Lotus japonicus*, *Pisum sativum* and *Arabidopsis thaliana*, for positional cloning of a legume symbiosis gene. *Theor. Appl. Genet.* 108, 442–449. doi:10.1007/s00122-003-1438-2

Tang, H., Krishnakumar, V., Bidwell, S., Rosen, B., Chan, A., Zhou, S., et al. (2014). An improved genome release (version Mti4.0) for the model legume *Medicago truncatula*. *BMC Genomics* 15:312. doi:10.1186/1471-2164-15-312

Thoquet, P., Ghérardi, M., Journet, E.-P., Kereszt, A., Ané, J.-M., Prosperi, J.-M., et al. (2002). The molecular genetic linkage map of the model legume *Medicago truncatula*: an essential tool for comparative legume genomics and the isolation of agronomically important genes. *BMC Plant Biol.* 2:1. doi:10.1186/1471-2229-2-1

Tominaga, A., Gondo, T., Akashi, R., Zheng, S.-H., Arima, S., and Suzuki, A. (2012). Quantitative trait locus analysis of symbiotic nitrogen fixation activity in the model legume *Lotus japonicus*. *J. Plant Res.* 125, 395–406. doi:10.1007/s10265-011-0459-1

Urbaniski, D. F., Malofeysky, A., Stougaard, J., and Andersen, S. U. (2012). Genome-wide LOREI retrotransposon mutagenesis and high-throughput insertion detection in *Lotus japonicus*. *Plant J.* 69, 731–741. doi:10.1111/j.1365-313X.2011.04827.x

Vandecasteele, C., Teulat-Merah, B., Morrière-Le Paven, M.-C., Leprince, O., Ly Vu, B., Viau, L., et al. (2011). Quantitative trait loci analysis reveals a correlation between the ratio of sucrose/raffinose family oligosaccharides and seed vigour in *Medicago truncatula*. *Plant Cell Environment* 34, 1473–1487. doi:10.1111/j.1365-3040.2011.02363.x

Vaughn, J. N., Nelson, R. L., Song, Q., Cregan, P. B., and Li, Z. (2014). The genetic architecture of seed composition in soybean is refined by genome-wide association scans across multiple populations. *G3* *Bethesda* 4, 2283–2294. doi:10.1534/g3.114.013343

Verdier, J., Torres-Jerez, I., Wang, M., Andriankaja, A., Allen, S. N., He, J., et al. (2013). Establishment of the *Lotus japonicus* Gene Expression Atlas (LjGEA) and its use to explore legume seed maturation. *Plant J. Cell Mol. Biol.* 74, 351–362. doi:10.1111/tpj.12119

Voisin, A.-S., Guéguen, J., Huyghe, C., Jeffroy, M.-H., Magrini, M.-B., Meynard, J.-M., et al. (2014). Legumes for feed, food, biomaterials and bioenergy in Europe: a review. *Agron. Sustain. Dev.* 34, 361–380. doi:10.1007/s13593-013-0189-y

Wakabayashi, T., Oh, H., Kawaguchi, M., Harada, K., Sato, S., Ikeda, H., et al. (2014). Polymorphisms of E1 and GIGANTEA in wild populations of *Lotus japonicus*. *J. Plant Res.* 127, 651–660. doi:10.1007/s10265-014-0649-8

Wen, Z., Tan, R., Yuan, J., Bales, C., Du, W., Zhang, S., et al. (2014). Genome-wide association mapping of quantitative resistance to sudden death syndrome in soybean. *BMC Genomics* 15:809. doi:10.1186/1471-2164-15-809

Yang, S., Gao, M., Xu, C., Gao, J., Deshpande, S., Lin, S., et al. (2008). Alfalfa benefits from *Medicago truncatula*: the RCT1 gene from *M. truncatula* confers broad-spectrum resistance to anthracnose in alfalfa. *Proc. Natl. Acad. Sci. U.S.A.* 105, 12164–12169. doi:10.1073/pnas.0802518105
Yano, K., Yoshida, S., Mueller, J., Singh, S., Banba, M., Vickers, K., et al. (2008). CYCLOPS, a mediator of symbiotic intracellular accommodation. Proc. Natl. Acad. Sci. U.S.A. 105, 20540–20545. doi: 10.1073/pnas.0806858105

Yoder, J. B., Stanton-Geddes, J., Zhou, P., Briskine, R., Young, N. D., and Tiffin, P. (2014). Genomic signature of adaptation to climate in Medicago truncatula. Genetics 196, 1263–1275. doi: 10.1534/genetics.113.159319

Young, N. D., Debelle, F., Oldroyd, G. E. D., Geurts, R., Cannon, S. B., Udvardi, M. K., et al. (2011). The medicago genome provides insight into the evolution of rhizobial symbioses. Nature 480, 520–524. doi: 10.1038/nature10625

Zhang, D., Song, H., Cheng, H., Hao, D., Wang, H., Kan, G., et al. (2014). The acid phosphatase-encoding gene GmACP1 contributes to soybean tolerance to low-phosphorus stress. PLoS Genet. 10:e1004061. doi: 10.1371/journal.pgen.1004061

Zhu, H., Choi, H. K., Cook, D. R., and Shoemaker, R. C. (2005). Bridging model and crop legumes through comparative genomics. Plant Physiol. 137, 1189–1196. doi: 10.1104/pp.104.058891

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2015 Gentzbittel, Andersen, Ben, Rickauer, Stougaard and Young. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.