MOLECULAR ECOLOGY

Molecular Ecology (2012) **21**, 4665–4668

NEWS AND VIEWS

PERSPECTIVE

Transcriptomic insights into mechanisms of symbiotic cooperation

ELLEN L. SIMMS and STEPHANIE S. PORTER
Department of Integrative Biology, University of California, 1001 Valley Life Science Building #3140, Berkeley, California, 94720-3140, USA

Intraspecific genetic variation can affect community structure and ecosystem processes (Bolnick *et al.* 2011). It can also influence phenotypic expression by genotypes within other species to produce genotype-by-genotype (G × G) interaction (Falconer & Mackay 1996). Evolution of one species drives correlated evolution of others when it causes G × G for fitness (Thompson 2005). However, the mechanisms by which species interact also influence evolutionary outcomes (Kummel & Salant 2006; Golubski & Klausmeier 2010; Akçay & Simms 2011; Grman *et al.* 2012). To identify genes and putative functional mechanisms underlying G × G interactions, Heath *et al.* (2012) analysed natural variation in the symbiotic transcriptome of the mutualistic nutritional symbiosis between a legume host *Medicago truncatula* and the facultative endosymbiotic rhizobium *Sinorhizobium meliloti*. Using twelve microarrays, the authors simultaneously measured host and symbiont gene expression in root nodules from four factorial pairings of host and symbiont genotypes that produced G × G in host fitness (Fig. 1, upper panel). Rhizobium gene expression was influenced by rhizobium and plant genotype and the G × G interaction (Fig. 1, lower panel), whereas plant gene expression was influenced primarily by plant genotype. The authors identified rhizobium genes that might contribute to G × G in host plant fitness. Heath *et al.* (2012) have moved beyond the constraints of single organism analysis towards a more realistic understanding of plants and bacteria as organisms inextricably linked with symbioses that affect even basic patterns of gene expression.

Received 27 April 2012; revised 14 June 2012; accepted 27 June 2012

Transcriptomics and symbiotic control

When one host exchanges goods or resources with multiple symbionts (Sachs *et al.* 2004), conflict over the rate of exchange (Bronstein 2001) will select for host regulation of the symbiont population and symbiont evasion or manipulation of host regulation (Simms & Taylor 2002; West *et al.* 2002; Sachs *et al.* 2004; Kummel & Salant 2006; Golubski & Klausmeier 2010). Theory predicts that, when hosts and symbionts negotiate the exchange rate, the evolutionary outcome depends on the exchange mechanisms, which affect which partners control the interaction (Akçay & Roughgarden 2007; Golubski & Klausmeier 2010; Akçay & Simms 2011; Grman *et al.* 2012). Transcriptomes provide large sets of quantitative traits that are free of the biases of traditional phenotypes. These might reveal the molecular bases of traits that affect control over exchange rates, which contribute to fitness G × G in symbioses.

In this issue of *Molecular Ecology* Heath *et al.* (2012) sought rhizobium genes available to co-evolutionary selection by using a false-discovery-rate (FDR) controlled two-way ANOVA to identify rhizobium genes expressed differently between plant genotypes, between rhizobium genotypes, and among the G × G combinations. They next focused on clusters of G × G interaction-responsive rhizobium genes that shared expression patterns. One co-expression cluster was positively correlated with plant benefit and enriched in symbiotic plasmid-borne genes, whereas two clusters were negatively correlated with plant benefit and enriched in chromosomal genes (Fig. 1). Thus, upregulation of symbiotic plasmid gene expression and downregulation of chromosomal gene expression are both available to co-evolutionary selection because both are genetically variable and correlated with partner benefit. Mechanisms of symbiotic interaction within the nodule have remained largely a black box; natural rhizobial transcriptional variation detected in this experiment represents a frontier for new functional discoveries. Only one G × G responsive plant gene was identified, possibly due to incomplete genome sampling and/or developmental heterogeneity whereby expression changes in some cells counterbalance opposing changes in others (Selimkhanov *et al.* 2012).

The effect of symbiosis on the relative fitness of both partners distinguishes antagonism from maladaptation (Table 1) and dictates the evolutionary trajectory (Sachs & Simms 2008). Thus, an exciting next step in symbiotic transcriptomics will be to associate transcription patterns with G × G interactions in the fitnesses of *both* partners. The authors' methods could also be extended to multiple G × G combinations from population-level samples (McGraw *et al.* 2011) of both partners to determine whether expression levels of particular host or symbiont genes are consistently associated with certain patterns of host and symbiont fitness. For example, if expression of a set of rhi-

Correspondence: Ellen L. Simms, Fax: 1-510-643-6264;
E-mail: esimms@berkeley.edu

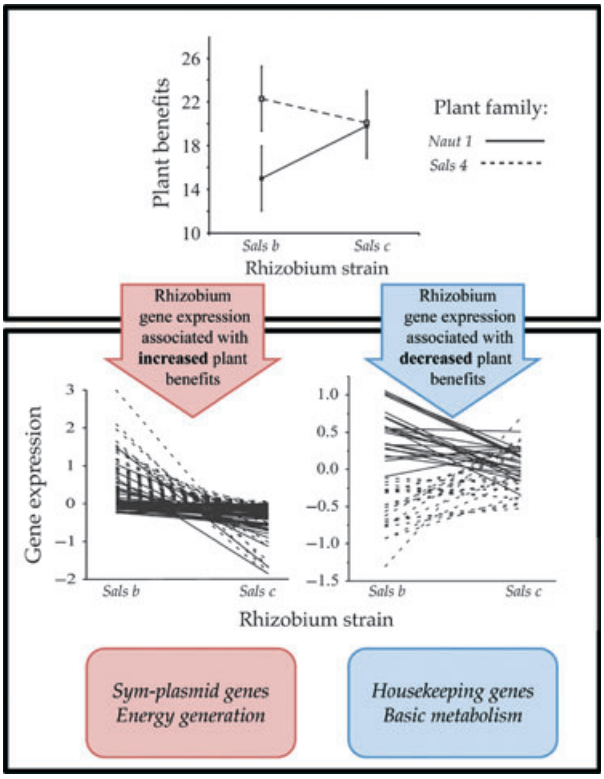

Fig. 1 Connecting natural variation in symbiotic gene expression to co-evolutionary phenotypes in the legume–rhizobium mutualism. Upper panel: plant fitness benefits depend on the plant genotype by rhizobium genotype (G × G) interaction. Lower panel: expression levels of rhizobium genes correlated with the G × G interaction in plant fitness. Figure courtesy of Heath *et al.* (2012).

zobium genes was associated with rhizobium benefit at host cost, these could underlie a mechanism to manipulate or evade host control. Thus, population variation in gene expression patterns could stimulate hypotheses about symbiotic control mechanisms. Heath *et al.* (2012) have taken an important step towards elucidating this variation, but their results should be interpreted cautiously due to the low number ($n = 4$) of G × G combinations. While it will be daunting to scale up such factorial symbiotic experiments, this is essential for predicting co-evolution between interacting partners and testing simplistic assumptions of current evolutionary models.

Including weakly responsive genes

Fitness components are massively polygenic and underlain by myriad genes of small effect (Lynch 2007; Reed *et al.* 2008; Ayroles *et al.* 2009). Further, the magnitude of differential expression does not necessarily indicate the importance of a gene to downstream phenotypic effects. Indeed, if a small change in a regulatory gene causes large responses in downstream genes, then an upstream gene

Table 1 Fitness relationships of host (legume) and symbiont (rhizobium) determine the nature and evolutionary trajectory of a symbiosis. In a mutualism, selection favours symbiosis because both host and symbiont experience higher fitness in symbiosis than when free-living. When partners are maladapted to each other, selection disfavours symbiosis because both experience lower fitness in symbiosis than when free-living. This pattern might arise when a rhizobium adapted to one host species or genotype partners with a novel host species or genotype. An individual that gains fitness at the expense of its partner is antagonistic (e.g. parasitic); selection on that individual favours symbiosis, whereas selection on its partner disfavours symbiosis. When an organism's fitness is unaffected by the interaction (i.e. relative symbiotic fitness = 1), selection on that organism neither favours nor disfavours symbiosis because the partner is commensal

	Relative symbiotic fitness*	Legume	
		>1	≤1
Rhizobium	>1	Mutualism	Antagonistic rhizobium
	≤1	Antagonistic legume	Maladaptation

*Fitness of each partner in symbiosis relative to its fitness when free-living.

that regulates more genes is more likely to be excluded by a FDR procedure (Almudevar *et al.* 2006). Separately analysing each gene identifies transcripts that exhibit large expression responses, but the FDR criterion discards information by excluding genes with weak responses.

The authors recaptured weakly responsive genes with a principal variance components analysis (PVCA). This procedure first uses principal components analysis (PCA) to identify a smaller number of axes (principal components, PCs) that together account for most transcription variance (Fig. 2). PCs incorporate genes with both small and large responses. Individual PCs are analysed by ANOVA to determine the proportion of variance (variance component) explained by each experimental factor (i.e. plant genotype, rhizobium genotype and G × G interaction). Summing the weighted values of variance components estimates the proportion of total transcriptional variance, caused by both small- and large-response genes, that was stimulated by that experimental treatment. For example, the plant genotype component of rhizobium transcriptional variance was summed across six PCs to estimate the proportion of rhizobium transcriptional variance caused by plant genotype.

The biotic environment (i.e. host genotype) was a key factor determining gene expression in symbionts. Plant genotype explained half as much variance in rhizobium gene expression as did rhizobium genotype. Further, transcription of rhizobium genotypes differed with host genotype: a small

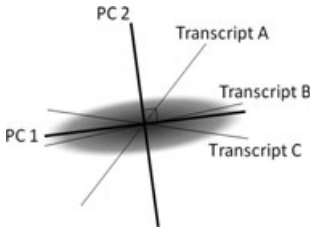

Fig. 2 The expression space of a hypothetical transcriptome spanned by three symbiont transcripts, labelled A, B and C, expressed in response to various combinations of host and symbiont genotypes. Each host–symbiont genotype combination is a point in this space. A large number of host–symbiont combinations would produce the depicted 'cloud' of points, which in three dimensions resembles an oblong bar of soap. The axes are not orthogonal because expression of the genes is correlated. In Heath *et al*. (2012), this space was spanned by hundreds of rhizobium transcripts and was occupied by 12 points representing the three replicates of each of the four pairwise combinations of two legume and two rhizobium genotypes. In the hypothetical example, transcript B exhibits the most variation in expression among host–symbiont combinations; so the first principal component (PC 1, the long axis of the soap bar) is nearly colinear with it. PC 2 must encompass the most remaining transcriptional variance yet also be perpendicular to PC 1 (i.e. align with the width of the soap bar). PC 3 (the shortest axis of the soap bar) would account for the remaining variance. In this example, PC 1 and PC 2 account for most of the transcriptional variance among symbiont combinations. Importantly, both axes incorporate some of the variance in expression of transcript A, which would have been excluded by a FDR procedure.

but significant G × G interaction. Thus, rhizobium gene expression might evolve in response to evolution in a host population. In contrast, plant gene expression did not differ with rhizobium genotype, and the G × G interaction component of plant transcription variance was negligible. Nearly a third of plant gene expression variance was explained by plant genotype and another third by the greenhouse block effect. Thus, abiotic environmental variation provoked stronger variation in plant transcription than did infection by different rhizobium strains. Do these gene expression patterns suggest the plant has more control over the interaction than do rhizobia?

Quantifying the effect of weakly responsive genes

The analytical framework of Heath *et al*. (2012) offers an opportunity to compare the relative importance of genes with significant vs. nonsignificant responses to experimental manipulations. The FDR-controlled ANOVAs identify individual genes exhibiting significant responses to host genotype, symbiont genotype or their interaction. The PVCA estimates the percentages of total transcriptional variance because of these same experimental factors, but does not distinguish the effects of genes that met the FDR criterion from those that did not. The relative importance to transcriptional variance of the latter genes could be estimated by performing PVCA on data sets from which genes that met the FDR criterion have been excluded. The variance components from this reduced data set could be compared to those calculated from the original PVCA, which encompasses genes with both significant and nonsignificant responses. This analysis could examine whether regulatory cascades initiated by weakly responsive genes influence ecologically important phenotypes (Almudevar *et al*. 2006).

References

Akçay E, Roughgarden J (2007) Negotiation of mutualism: rhizobia and legumes. *Proceedings of the Royal Society B-Biological Sciences*, **274**, 25–32.

Akçay E, Simms EL (2011) Negotiation, sanctions, and context dependency in the legume-rhizobium mutualism. *American Naturalist*, **178**, 1–14.

Almudevar A, Klebanov LB, Qiu X, Salzman P, Yakovlev AY (2006) Utility of correlation measures in analysis of gene expression. *NeuroRX*, **3**, 384–395.

Ayroles JF, Carbone MA, Stone EA *et al*. (2009) Systems genetics of complex traits in *Drosophila melanogaster*. *Nature Genetics*, **41**, 299–307.

Bolnick DI, Amarasekare P, Araúo MS *et al*. (2011) Why intraspecific trait variation matters in community ecology. *Trends in Ecology & Evolution*, **26**, 183–192.

Bronstein JL (2001) The exploitation of mutualisms. *Ecology Letters*, **4**, 277–287.

Falconer DS, Mackay TFC (1996) Introduction to Quantitative Genetics. Longman Group Ltd., Essex, UK.

Golubski AJ, Klausmeier CA (2010) Control in mutualisms: combined implications of partner choice and bargaining roles. *Journal of Theoretical Biology*, **267**, 535–545.

Grman E, Robinson TMP, Klausmeier CA (2012) Ecological specialization and trade affect the outcome of negotiations in mutualism. *The American Naturalist*, **179**, 567–581.

Heath KD, Burke PV, Stinchcombe JR (2012) Coevolutionary genetic variation in the legume-rhizobium transcriptome. *Molecular Ecology*, **21**, 4735–4747.

Kummel M, Salant SW (2006) The economics of mutualisms: optimal utilization of mycorrhizal mutualistic partners by plants. *Ecology*, **87**, 892–902.

Lynch M (2007) The Origins of Genome Architecture. Sinauer Associates, Sunderland, MA.

McGraw EA, Ye YH, Foley B *et al*. (2011) High-dimensional variance partitioning reveals the modular genetic basis of adaptive divergence in gene expression during reproductive character displacement. *Evolution*, **65**, 3126–3137.

Reed DR, Lawler MP, Tordoff MG (2008) Reduced body weight is a common effect of gene knockout in mice. *Bmc Genetics*, **9**, 6.

Sachs JL, Simms EL (2008) The origins of uncooperative rhizobia. *Oikos*, **117**, 961–966.

Sachs JL, Mueller UG, Wilcox TP, Bull JJ (2004) The evolution of cooperation. *Quarterly Review of Biology*, **79**, 135–160.

Selimkhanov J, Hasty J, Tsimring LS (2012) Recent advances in single-cell studies of gene regulation. *Current Opinion in Biotechnology*, **23**, 34–40.

Simms EL, Taylor DL (2002) Partner choice in nitrogen-fixation mutualisms of legumes and rhizobia. *Integrative and Comparative Biology*, **42**, 369–380.

Thompson JN (2005) The Geographic Mosaic of Coevolution, University of Chicago Press. Chicago.
West SA, Kiers ET, Pen I, Denison RF (2002) Sanctions and mutualism stability: when should less beneficial mutualists be tolerated? *Journal of Evolutionary Biology*, **15**, 830–837.

E.L.S. studies the ecology and evolution of biotic interactions involving plants. Her recent work focuses on legume-rhizobium interactions and their role in community response to environmental change. S.P. is an evolutionary ecologist who studies adaptation in symbiotic mutualisms, with a focus on plant-microbe symbioses across heterogeneous environments and the role of symbiosis in biological invasions.

doi: 10.1111/j.1365-294X.2012.05719.x