Dual Affymetrix GeneChip® analysis of the perennial ryegrass-endophyte symbiosis

AgResearch Grasslands, Private Bag 11008, Palmerston North, New Zealand
* linda.johnson@agresearch.co.nz

Abstract
Grass associations with *Epichloë*/*Neotyphodium* endophytes display enhanced fitness as well as prolonged field persistence over their endophyte free equivalents. To gain a comprehensive understanding of the complex biological interactions that occur between a plant host and fungal symbiont, a transcriptomics approach using custom designed Affymetrix GeneChip®s was employed. We are currently comparing and analysing symbiotic interactions of perennial ryegrass (*Lolium perenne*) with endophytes *N. lolii* and *E. festucae* as well as comparing it against endophyte-free perennial ryegrass. Both *N. lolii* and *E. festucae* have been grown in culture in order to compare *in planta* versus *in vitro* gene expression. Additionally, targeted gene replacements in *E. festucae* have been performed and a comparative analysis of the knock-outs with wild-type *E. festucae* infected plants is in progress. These comparative analyses have revealed changes in gene expression which may lead to the identification of gene pathways/networks and the roles of these genes in symbiosis.

Keywords: *Neotyphodium*, *Epichloë*, transcriptomics, symbiosis

Table 1 Summary of comparisons performed between target treatments. *E.f.*, *E. festucae* Fl1; *L.p.*, *L. perenne*; *N.l.*, *N. lolii* Lp19.

<table>
<thead>
<tr>
<th>Treatment 1</th>
<th>Treatment 2</th>
<th>Objective</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 <em>E.f.</em> in <em>L.p.</em> leaf sheath</td>
<td><em>E.f.</em> in liquid culture (PDB)</td>
<td><em>E.f.</em> genes expressed in symbiosis</td>
<td>This paper</td>
</tr>
<tr>
<td>2 <em>N.l.</em> in <em>L.p.</em> leaf sheath</td>
<td><em>N.l.</em> in liquid culture (PDB)</td>
<td><em>N.l.</em> genes expressed in symbiosis</td>
<td>This paper</td>
</tr>
<tr>
<td>3 <em>E.f.</em> in <em>L.p.</em> leaf sheath</td>
<td>Endophyte free <em>L.p.</em> leaf sheath</td>
<td>Effects of <em>E.f.</em> endophyte infection on <em>L.p.</em> genes</td>
<td>This paper</td>
</tr>
<tr>
<td>4 <em>N.l.</em> in <em>L.p.</em> leaf sheath</td>
<td>Endophyte free <em>L.p.</em> leaf sheath</td>
<td>Effects of <em>N.l.</em> endophyte infection on <em>L.p.</em> genes</td>
<td>This paper</td>
</tr>
<tr>
<td>5 <em>E.f.</em> in <em>L.p.</em> leaf sheath</td>
<td><em>N.l.</em> in <em>L.p.</em> leaf sheath</td>
<td>Differences between endophyte strains <em>in planta</em></td>
<td>This paper</td>
</tr>
<tr>
<td>6 <em>E.f.</em> in liquid culture (PDB)</td>
<td><em>N.l.</em> in liquid culture (PDB)</td>
<td>Differences between endophyte strains in culture</td>
<td>This paper</td>
</tr>
<tr>
<td>7 <em>E.f.</em> ΔSidF in <em>L.p.</em> leaf sheath</td>
<td><em>E.f.</em> in <em>L.p.</em> leaf sheath</td>
<td>Effects of endophyte siderophore deletion on fungal and plant genes</td>
<td>Johnson et al. 2007a</td>
</tr>
<tr>
<td>8 <em>E.f.</em> ΔNRPS1 in <em>L.p.</em> leaf sheath</td>
<td><em>E.f.</em> in <em>L.p.</em> leaf sheath</td>
<td>Effects of deletion of unknown endophyte 2° metabolite on fungal and plant genes</td>
<td>This paper</td>
</tr>
<tr>
<td>9 <em>E.f.</em> ΔacyA in <em>L.p.</em> pseudostem</td>
<td><em>E.f.</em> in <em>L.p.</em> pseudostem</td>
<td>Effects of deletion of endophyte cAMP signalling on fungal and plant genes</td>
<td>Voisey et al. 2007</td>
</tr>
<tr>
<td>10 <em>E.f.</em> ΔNC25 in <em>L.p.</em> pseudostem</td>
<td><em>E.f.</em> in <em>L.p.</em> pseudostem</td>
<td>Effects of endophyte NC25 deletion (function unknown) on fungal and plant genes</td>
<td>Johnson et al. 2007b</td>
</tr>
</tbody>
</table>
Table 2 Counts and brief descriptions of three fungal probe set types.

<table>
<thead>
<tr>
<th>Classification</th>
<th>Number of probe sets</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Type (i)</td>
<td>1201</td>
<td>Expressed in vitro only</td>
</tr>
<tr>
<td>Type (ii)</td>
<td>1421</td>
<td>Expressed in vitro and in planta</td>
</tr>
<tr>
<td>Type (iii)</td>
<td>225</td>
<td>Only expressed in planta (symbiotic)</td>
</tr>
<tr>
<td>Total</td>
<td>2847</td>
<td></td>
</tr>
</tbody>
</table>

Introduction

*Epichloë* and *Neotyphodium* endophytes provide various benefits to their plant hosts, primarily protection from biotic and abiotic stresses (reviewed in Schardl & Phillips 1997; Schardl et al. 2004). Plant physiological changes induced by the endophyte enable the host to quickly adapt to a variety of challenging circumstances giving infected grasses a competitive advantage. Four well-characterised classes of endophyte alkaloids, peramine, ergot alkaloids, loline, and lolitremes have been shown to contribute to grass fitness (Siegel & Bush 1996; Lane et al. 2000; Wilkinson et al. 2000; Tanaka et al. 2005; Ball et al. 2006; Panaccione et al. 2006) and it is hypothesised that other fungal bioactive metabolites that have yet to be characterised will also endow the grass host with specific fitness improvements. Additionally, the host also appears to have a role in the regulation of fungal products since fungal genes involved in alkaloid production (as well as the metabolites themselves) are up-regulated in planta (Panaccione et al. 2001; Spiering et al. 2005; Tanaka et al. 2005; Young et al. 2006; Fleetwood et al. 2007) in addition to some extracellular enzymes (Reddy et al. 1996; Moy et al. 2002) and other fungal genes of unknown function (Johnson et al. 2003; Johnson et al. 2006). Recently, the deletion of two endophyte genes involved in the generation of reactive oxygen species (Takemoto et al. 2006; Tanaka et al. 2006) and the deletion of an endophyte gene encoding an extracellular siderophore (Johnson et al. 2007a) showed that these genes are essential for maintenance of endophyte-grass associations. However, our understanding about these symbiotic interactions is still extremely limited. Genes with a symbiotic function are still waiting to be discovered as are those that are important for the regulation and production of endophyte bioactive metabolites. With the advent of the functional genomics era, we are able to apply techniques such as transcriptomics to comprehensively analyse global changes in gene expression. Here, to further unravel the complex molecular interplay between endophytes and their grass hosts we performed a transcriptomic study using an Affymetrix dual genome (*N. lolii/L. perenne*) GeneChip® to analyse gene expression from perennial ryegrass in association with either a *Neotyphodium* species (*N. lolii*) or an *Epichloë* species (*E. festucae*), as well as from these endophytes grown under *in vitro* conditions. Additionally, we have hybridised our GeneChip®s with perennial ryegrass plants associated with mutant endophyte strains that contain a single gene deletion of interest. These analyses have enabled us to identify candidate ‘symbiotic’ genes; an important subset of fungal or ryegrass genes that are only induced during the symbiotic stage and are therefore key candidates for further investigation.

Methods

Fungal growth conditions

Fungal mycelia from *E. festucae* (FL1) and *N. lolii* (Lp19) for RNA extraction and subsequent labelling and hybridisation was harvested from potato dextrose broth (PDB) cultures grown at 22°C for 3 or 5 days for FL1 and Lp19 respectively. Under these conditions a similar amount of fungal biomass was obtained from the two strains.

Plant inoculations and growth conditions

Isogenic plant lines of perennial ryegrass infected with FL1, Lp19, ΔSidF, ΔNRPS1, ΔacyA and ΔNC25 strains were obtained by micropropagation of auxiliary buds from endophyte-free perennial ryegrass tillers (G1057, cultivar ‘Nui’). Plants destined for RNA preparation and subsequent labelling and GeneChip® hybridisation were grown for 14 weeks in replicates of six under climate chamber conditions (HortResearch, Palmerston North, New Zealand) of 20°C day/10°C night, with a 14 h day length (653 μmoles). Only three biological replicates were chosen for RNA extraction, labelling and hybridisation.

Development of a Dual Genome (*N. lolii/L. perenne*) Symbiosis GeneChip®

Custom designed DNA microarrays were developed using NimbleGen technology (NimbleGen Systems, Inc.) through Affymetrix (California). Methodology on GeneChip® development and design is described by Voisey et al. (2007). To summarise, sequences tiled on the chip (with a maximum capacity of 12 000 open reading frames (ORFs) were derived primarily from three sources; ESTs from *N. lolii* (Lp19) grown in either liquid PDB or liquid minimal media (MM), ESTs from six suppressive subtractive hybridisation libraries from perennial ryegrass in association with *N. lolii* (Lp19), or from selected *L. perenne* ESTs. A smaller number of ORFs were obtained from endophyte NCBI sequence submissions, in-house endophyte sequences (from degenerate PCR libraries, lambda genomic Lp19 DNA libraries) and in-house Lp19 BAC sequences. The proportion of probe sets tiled on the chip was estimated to be approximately 31% from the fungus and 55% from *L. perenne* (Voisey et al. 2007).

Table 3 Criteria used to define fungal probe sets as one of four possible fungal probe set types. PMA calls were used to determine whether a probe set was expressed on a particular chip subtype or not.

<table>
<thead>
<tr>
<th>Chip subtype</th>
<th>(i)</th>
<th>(ii)</th>
<th>(iii)</th>
<th>(iv)</th>
</tr>
</thead>
<tbody>
<tr>
<td>(a) Endophyte only</td>
<td>√</td>
<td>√</td>
<td></td>
<td>√</td>
</tr>
<tr>
<td>(b) Ryegrass only</td>
<td></td>
<td></td>
<td>√</td>
<td></td>
</tr>
<tr>
<td>(c) Endophyte and Ryegrass</td>
<td>√</td>
<td>√</td>
<td></td>
<td>√</td>
</tr>
</tbody>
</table>
Preliminary microarray analysis
A preliminary analysis of the microarray data has been performed. The data were normalised using Robust Multichip Analysis (RMA; Irizarry et al. 2003) with sequences that were deduced to be derived from the two species on the chip, _L. perenne_ and _N. lolii_, analysed separately and treated independently (for further details see Voisey et al. 2007). A modified t-test was then performed on each probe set using the “limma” package in the statistical package Bioconductor (Smyth 2005). P-values were adjusted for multiple testing using the Benjamini and Hochberg correction (Benjamini & Hochberg 1995).

Additionally, ‘Present’, ‘Marginal’ and ‘Absent’ expression calls were generated for each probe set using the “masScalls” procedure in the “affy” BioConductor package (Gautier et al. 2004). This procedure uses the Wilcoxon signed rank test to classify the genes as ‘Present’ or otherwise. We are currently exploring the optimum normalisation procedure for analysis of a dual plant/fungal genome chip.

RNA extraction, labelling, hybridisation and scanning
Total RNA was partially extracted from samples using TRIZOL Reagent (Invitrogen) up to the aqueous phase stage. This was then transferred to a fresh tube where an equal volume of 70% ethanol was added before loading onto a Qiagen RNeasy mini column for column purification of total RNA using the manufacturer’s instructions (Qiagen). The RNA was analysed for quality using an Agilent Bioanalyser and labelled using the Affymetrix GeneChip® IVT Labeling Kit. Analysis of RNA quality, labelling, chip hybridisations, washes and GeneChip® scanning followed recommended procedures and were conducted at the Centre for Genomics and Proteomics, School of Biological Sciences, University of Auckland, New Zealand.

Results and Discussion
Biological material compared in GeneChip® hybridisations
Table 1 displays the biological material which was subjected to Affymetrix GeneChip® analysis and lists the comparisons made between treatments to enable identification of subsets of differentially expressed genes of biological interest. Genes were termed up- or down-regulated in Treatment 1 when they had a False Discovery Rate (FDR) ≤ 0.05 and a log_2_ ratio >1.0 or < -1.0 respectively.

Ten comparisons in total are currently being analysed and the overall aims of these experiments are: (1) to identify plant and fungal genes differentially expressed during symbiosis and to discover candidate ‘symbiotic’ genes; (2) elucidate gene expression differences between _Epichloë_ and _Neotyphodium_ endophytes (using strains FL1 and Lp19 respectively); (3) determine the effects of specific fungal gene deletions on all plant and fungal genes tiled on the chip. The results presented here focus on the identification of fungal genes that may have a functional role in symbiosis.

Experimental determination of fungal probe sets using PMA calls
Present, Marginal and Absent (PMA) calls provide information as to whether a probe set (equivalent to approximately eleven 25 bp oligomers designed along the length of one sequence tiled on the chip) has a detectable expression value (one that is present, marginal or absent). Based on the samples hybridised to individual chips (see Table 1 for descriptions), we defined four types of fungal probe sets that are expressed on the chips. Type (i) fungal probe sets are those that are expressed _in vitro_ and not _in planta_; type (ii) are fungal probe sets that are expressed _in vitro_ as well as _in planta_; and type (iii) fungal probe sets are those that are expressed _in planta_ but not _in vitro_ and are therefore categorised as symbiotic probe sets (Table 2). A further category, type (iv), consists of probe sets to which RNA from both fungal and ryegrass sources bind (cross-hybridise). Many type (iv) probe sets appear to represent highly conserved genes (such as beta-tubulin and ubiquitin), and will therefore not be discussed further in this paper. To allocate the probe sets into species categories, three chip subtypes were first defined based on which type of RNA was hybridised to a slide: (a) endophyte only; (b) ryegrass only; or (c) fungal and ryegrass RNA. Probe sets were then allocated to probe set category types (i) – (iv) based on which chip subtypes they were expressed on, as shown in Table 3.

For example, type (i) probe sets are probe sets which are expressed on the endophyte only chip subtype (subtype a) but not on either the Ryegrass only (subtype b) or Endophyte and Ryegrass chip subtype (subtype c).

Using these criteria we determined that the majority of fungal probe sets tiled on the chip were expressed under both _in planta_ and _in vitro_ conditions (Table 2), although a significant number were expressed only _in vitro_. Additionally, a set of 225 probe sets were found to be only switched on _in planta_ and hence are flagged as symbiotic probe sets. The subset of genes which hybridise to the symbiotic probe sets warrant further investigation as they are highly likely to play a role in symbiosis. It will also be interesting to determine whether these genes are fungal genes which are only expressed _in planta_, or whether they are ryegrass genes which are only expressed when the plant is infected with the endophyte.

Identification of endophyte genes differentially expressed in the symbiosis
In Table 1, comparisons 1 and 2 aim to compare genes expressed from Treatment 1 containing either _E. festucae_ FL1 or _N. lolii_ Lp19 grown _in planta_ (_L. perenne_ leaf sheaths) versus expression (of each respective strain) _in vitro_ (grown in liquid PDB) from Treatment 2. Based on FDR and log, ratio cut offs, we have created candidate gene lists that differentiate expression characteristics of the two fungal endophyte strains, Lp19 and FL1. Within each endophyte strain, probe sets have been classified as either up-regulated _in vitro_ (compared to _in planta_), up-regulated _in planta_ (compared with expression _in vitro_), or not differentially expressed _in planta_ and _in vitro_ (Fig. 1). The latter category comprises both probe sets which had significantly different mean intensities between the two categories (FDR ≤ 0.05) but for which this difference was minimal (absolute log_2_ ratio between 1 and -1), and also includes those genes which were not differentially expressed _in planta_ _versus_ _in vitro_ (FDR>0.05). The number of probe sets _up-regulated in vitro_ is very similar in both Lp19 (865) and FL1 (821). These numbers are also of the same order of magnitude as the number of probe sets estimated to be only expressed in culture via the PMA method (1201). Both methods suggest that a considerable number of genes are either only expressed _in vitro_, or are expressed at a much higher level when the fungus is grown _in planta_ as opposed to _in planta_. Given how different the conditions are for the fungus in culture versus _in planta_, it is not surprising that a different suite of genes are switched on. There are also approximately 300 more probe sets _up-regulated in planta_ for Lp19 (1173) compared to FL1 (821). Although this difference may be biologically significant, it may also simply be a reflection of the fact that the probe sets tiled on the chip originate mostly from Lp19.
Figure 1 Fungal probe set subgroups (as defined by FDR and log₂-ratio cut-offs). Up-regulated probe sets were defined as those with an FDR ≤0.05 in conjunction with log₂ cut offs. The log₂ ratio cut off for in vitro probe sets was < -1 and for in planta a log₂ > 1 was applied. The non-differentially expressed probe sets were defined by an absolute log₂ ratio between 1 and -1 and with an FDR ≤0.05 and included probe sets with an FDR ≥0.05 without a log₂ cut off applied.

The most interesting subset of genes are those that are upregulated in planta. A large portion of these genes display very high log₂ ratios (log₂ = 9.6). The most abundantly expressed gene from FL1-infected perennial ryegrass is Nc25; a novel fungal gene previously shown to be highly expressed in different plant-fungal associations, but absent in cultures of Neotyphodium coenophialum (Johnson et al. 2003; Johnson et al. 2007b). Other genes from within this high log₂ category are lolitrem biosynthetic pathway genes. These genes are expressed abundantly in planta and at low levels or not at all in culture (Young et al. 2006). In addition, ergovaline biosynthetic pathway genes are also highly expressed as expected (Fleetwood et al. 2007). Most exciting are the number of highly expressed fungal genes that have no annotations (based on BLAST homology to sequences in public domains and InterProScan searches as described in Voisey et al. 2007), indicating that they are novel. Future investigation of these genes should yield insights into the molecular regulation of grass-endophyte symbioses. However, it must be noted that the number of genes estimated to be up-regulated in planta may be presently overestimated. This is because the high number of fungal probe sets which are not expressed in planta (see Table 2) may distort the RMA normalisation, as this is a rank-based method and so requires similar numbers of genes to be expressed in both treatments to work successfully. Work is currently underway to address this problem, with likely solutions being the use of the control putative “housekeeping” genes tiled on the chip (see Voisey et al. 2007). Note that the same problem does not apply to the genes up-regulated in culture.

Comparative analyses of fungal genes up-regulated in planta versus those in culture clearly show that the majority of annotated genes up-regulated in planta are those involved in secondary metabolism, whereas the types of genes up-regulated in culture are housekeeping genes such as those encoding ribosomal proteins, RNA polymerase II, ATP synthase, cytochrome c etc.

Continuing research/analyses
All of our tiled genes have been annotated with Gene Ontology (GO) terms (The Gene Ontology Consortium 2000) to aid in functional classification, as well as to support identification of enriched functional categories present in differentially expressed gene lists derived from each of the comparisons listed in Table 1. The GO analyses will provide further clues on biological processes and pathways which are up- or down-regulated in each of our conditions tested. Another interesting avenue for further exploration will be the genes of unknown function represented on the GeneChip®. We hope to establish putative functions for many of these genes using co-expression clustering across all available samples. By applying a ‘guilt by association’ paradigm, we can hypothesise that unknown genes which are co-expressed with characterised genes may share the same transcriptional regulatory inputs and therefore may share similar functions.

In summary, we have generated a number of differentially expressed gene lists from the different conditions tested (see Table 1). We have used two complementary methods to identify fungal genes that are highly likely to play a role in some aspect of the symbiosis. Reported here is the discovery of 225 symbiotic plant and/or fungal genes where expression of these was only detected on chips hybridised with endophyte-infected perennial ryegrass. We have also identified approximately 800 or 1100 endophyte genes up-regulated in planta versus those in culture clearly show that the majority of annotated genes up-regulated in planta are those involved in secondary metabolism, whereas the types of genes up-regulated in culture are housekeeping genes such as those encoding ribosomal proteins, RNA polymerase II, ATP synthase, cytochrome c etc.

ACKNOWLEDGEMENTS
We acknowledge Catherine Tootill for excellent technical assistance and funding support from the New Zealand Foundation of Research Science and Technology under contract AGRX0204 and PROJ-1033_ECOS_AGR and also C10X0203.

REFERENCES
Ball, O.J.P.; Coudron, T.A.; Tapper, B.A.; Davies, E.; Trently, D.; Bush, L.P.; Gwinn, K.D.; Popay, A.J. 2006. Importance of host plant species, Neotyphodium endophyte isolate, and


