Analysis of expressed sequence tags derived from the endophytic fungus *Neotyphodium lolii* grown in vitro and in association with its host plant perennial ryegrass

R. JOHNSON¹, A. KHAN², C. VOISEY¹, S. BASSETT¹, C. GABORIT¹, L. JOHNSON³, M. CHRISTENSEN¹, A. MCCULLOCH² and G. BRYAN¹.

¹ Forage Biotechnology, AgResearch, Private Bag 11008, Palmerston North, New Zealand.
² Bioinformatics, Mathematics and Statistics, AgResearch, Private Bag 50034, Mosgiel, New Zealand
³ Information Services, AgResearch, Private Bag 50034, Mosgiel, New Zealand
richard.johnson@agresearch.co.nz

Abstract

As a first step towards a functional genomics approach to gain a greater understanding of this important symbiosis, we have generated, sequenced and analysed two EST libraries from cultures of *N. lolii* and six in planta subtracted EST libraries enriched for differentially expressed genes. A total of 12871 ESTs were sequenced which, after filtering for quality, clustered into 1066 contigs and 3230 singletons to give a set of 4296 unique sequences or unigenes. BLASTX analysis revealed that 60% of fungal sequences derived from cultures were of unknown function with a sub-set of these corresponding to orphans. For the in planta-derived ESTs, most of the sequences with homologs in the public databases (98%) were of ryegrass origin. Comparisons made against fully sequenced genomes revealed that most fungal ESTs were homologous to genes present in both pathogenic and non-pathogenic ascomycete filamentous fungi, whereas the subtracted libraries comprised mostly plant genes. A range of sequences having significant homology to demonstrated pathogenicity/virulence genes in other fungal pathosystems were also identified, as well as some ESTs with proven roles in endophyte secondary metabolism.

Keywords: ESTs, cDNA, *Neotyphodium lolii*, *Lolium perenne*, symbiosis, mutualism, suppression subtractive hybridisation

Introduction

Forage grasses belonging to the sub-family pooidae, including several important forage and turf species, often harbour endophytic fungi belonging to the genus *Neotyphodium* and *Epichloë* (Clavicipitaceae, Ascomycota). These endophytes live entirely within the intercellular spaces of their grass hosts and infection is symptomless, with the endophyte relying entirely on the host plant for dissemination via the seed or through vegetative structures (reviewed by Schardl et al. 2004).

The association is mutually beneficial since the endophyte confers a number of biotic and abiotic advantages to the host, including enhanced plant growth, protection from certain mammalian and insect herbivores, enhanced resistance to nematodes, resistance to some fungal pathogens and in some associations, enhanced drought tolerance (Johnson et al. 2003). Some of these benefits are due to the production of fungal secondary metabolites such as peramine, loline, ergopeptide and indole diterpene alkaloids (Tanaka et al. 2005; Spiering et al. 2005; Panaccione et al. 2001; Young et al. 2006), some of which also cause associated mammalian toxicosis. Evidence has also accumulated showing that the host plant has a significant effect on the regulation of fungal secondary metabolites (Lane et al. 2000) and more recently it has been shown that the expression of fungal genes involved in alkaloid production are up-regulated in planta (Tanaka et al. 2005; Young et al. 2005).

Apart from the characterised role of the above mentioned fungal secondary metabolites during symbiosis, many of the other observed endophyte effects on their host plants have not been elucidated. Whilst genomic resources for fungal researchers have significantly advanced over the past few years, particularly in respect to ESTs, these have generally focused on pathogenic, model saprophytic or important industrial fungi. In this paper we describe a functional genomics approach to dissect the molecular mechanisms underlying the symbiosis between endophytic fungi and their grass hosts, in which we generated, sequenced and analysed two EST libraries from cultures of *N. lolii* and six in planta subtracted EST libraries enriched for differentially expressed genes.

Materials and Methods

Endophyte strains and plant infection

An asexual wild type *N. lolii* strain, Lp19 (Christensen et al. 1993) isolated from Nui perennial ryegrass was chosen for this study because it synthesises three of the four important known symbiosis-associated secondary metabolites (peramine, ergovaline and lolitrem), has a relatively small genome size (~35 Mb), is stable in culture and is typical of endophytes isolated from perennial ryegrass pastures. Isogenic ryegrass plants infected (G1056) or uninfected (G1057) with *N. lolii* strain Lp19 were obtained as previously described (Tanaka et al. 2005). Plants of G1056 and G1057 plants were grown under identical conditions under glass until harvest.

Growth of endophyte strains in vitro

To generate ESTs from endophytes grown in a complete medium, *N. lolii* strain Lp19 was grown in 50 ml potato dextrose broth and incubated for 10 days at 22°C. Mycelia for RNA extraction were harvested under vacuum using a sterile Buchner funnel containing two layers of Whatman 3 MM paper. To generate ESTs from endophyte grown in a minimal medium, mycelia from cultures initially grown in complete medium for 14 days were harvested as described above. Two grams of mycelia was used to inoculate 50 ml of Blankenship medium (Blankenship et al. 2001) and the cultures incubated for a further 19 days at 22°C. Mycelia for RNA extraction were harvested as for complete medium.

Preparation of in vitro grown cDNA libraries

Polyadenylated mRNA was purified from total RNA using standard procedures. cDNA libraries were created using the SMART cDNA Library Construction Kit (Clontech) and T3/7primer vector (Clontech) according to the manufacturers instructions. Two libraries were constructed, one from minimal media (NL19MM) and one from defined media (NL19EPD). ESTs were sequenced using Big Dye 3.1 (ABI) cycle sequencing reagents on an ABI 3100 automated sequencer.
Table 1 Numbers of unisequences that have homology to known or unknown proteins, and those that are orphans. The unknowns category includes those unisequences with hits to hypothetical/predicted proteins, and those with no hits (orphans).

<table>
<thead>
<tr>
<th>Contig build</th>
<th>Homology to known proteins</th>
<th>Homology to unknown proteins</th>
<th>Orphans</th>
</tr>
</thead>
<tbody>
<tr>
<td>CS32</td>
<td>695 (34.6%)</td>
<td>1,316 (65.4%)</td>
<td>1,224 (60.9%)</td>
</tr>
<tr>
<td>CS36</td>
<td>1,520 (66.5%)</td>
<td>765 (33.5%)</td>
<td>318 (13.9%)</td>
</tr>
</tbody>
</table>

Endophyte infected and endophyte free perennial ryegrass (together designated CS36). A total of 5493 and 3896 ESTs were sequenced for CS32 and CS36, respectively. For library CS32, ESTs assembled into 412 contigs and 1599 singletons to provide a set of 2011 unisequences with an average length of 716 base pairs. For library CS36, ESTs assembled into 654 contigs and 1631 singletons to provide a set of 2285 unisequences with an average length of 377 base pairs. Overlap of expression between the two in vitro fungal derived libraries that comprise CS32 was determined on the basis of contig membership (Fig. 1A). For the subtracted SSH libraries we compared the overlap between the combined up-regulated and down-regulated libraries (Fig. 1B). In both cases the majority of contigs or singletons showed no overlap indicating that most unisequences were library specific.

EST comparison against fungal and plant databases
We analysed the contigs in CS32 and CS36 to determine what proportion of unisequences had significant homology to proteins of known or unknown function, as well as determining the percentage of orphan sequences (Table 1). Contig builds were compared by BLASTX analysis against the NCBI non redundant (nr) protein database, using an e-value cut-off of 1e-5, and the top five hits for each contig were examined. Orphans were confirmed using additional TBLASTX searches against ESTs and ORFs from 25 fungal and 11 plant species.

Functional classification of CS32 and CS36
We have annotated each EST in CS32 and CS36 with one or multiple functional categories which define the function of the gene product. Functional categories were selected from the MIPS Functional Catalogue (FunCat) (Ruepp et al. 2004) and are represented as pie charts for CS32 and CS36, respectively (Figure 2A and 2B). BLASTX analysis revealed that 60% of EST sequences from cultures of N. lolii were of unknown function with a sub-set of these corresponding to orphans. Comparisons made against fully sequenced genomes revealed that most fungal ESTs were homologous to genes present in both pathogenic and non-pathogenic ascomycete filamentous fungi, whereas the subtracted libraries comprised mostly plant genes.

Discussion
This study has provided significant EST resources for N. lolii grown in vitro and this has been of critical importance in aiding our proteomics approach in which we have identified proteins differentially expressed during this symbiosis using the same infected and uninfected isogenic plant lines reported here (Johnson et al. 2007).

We have also investigated the molecular genetics of the symbiosis between Neotyphodium lolii and its host plant Lolium perenne. Unlike many fungal pathogens, where fungal biomass can accumulate to relatively high levels in relation to their hosts, endophytic fungi of the Neotyphodium/Epichloë complex primarily colonise the meristematic parts of their grass hosts and fungal biomass has been shown to be concomitantly low; less that 0.5% at the level of DNA (Panaccione et al. 2001; Johnson et al. 2003). This presents a technical challenge to identify N. lolii ESTs that are expressed in planta, which is why we used suppression subtractive hybridisation to enrich specifically for ESTs expressed during the symbiosis. Analysis of the SSH libraries identified several fungal genes that have a known role in the symbiosis, for example lolitrem biosynthesis (Young et al. 2006), β-1, 6-glucanase (Moy et al. 2002), a proteinase (Reddy et al. 1996), a chitinase (Li et al. 2004), as well as others with
no obvious homology to other genes. Johnson et al. (2003) also identified two fungal genes (Nc12 and Nc25), from the tall fescue/Neotyphodium association, in common with this study.

To obtain accurate data on the expression of these and other candidate genes we have developed the first dual genome Affymetrix chip (Voicey et al. 2007) comprising all non-redundant unisequences (both CS32 and CS36) generated in the current study. Confirmation of which genes are differentially expressed during symbiosis and subsequent functional analysis, by gene disruption or RNAi, will aid in elucidating their possible roles in the symbiosis.

ACKNOWLEDGEMENTS
Funding has been provided by the Foundation of Research Science and Technology (C10X0203).

REFERENCES
Christensen, M.J.; Leuchtmann, A.; Rowan, D.D.; Tapper, B.A.


Reddy, P.V.; Lam, C.K.; Belanger, F.C. 1996. Mutantlich fungal endophytes express a proteinase that is homologous to proteases suspected to be important in fungal pathogenicity. *Plant Physiology* 111: 1209-1218.


