Generation of *Phaseolus vulgaris* ESTs and investigation of their regulation upon *Uromyces appendiculatus* infection

Sandra Thibivilliers¹, Trupti Joshi², Kimberly B Campbell³, Brian Scheffler⁴, Dong Xu², Bret Cooper³, Henry T Nguyen¹ and Gary Stacey*¹

Address: ¹National Center for Soybean Biotechnology, Center for Sustainable Energy, Divisions of Plant Sciences and Biochemistry, University of Missouri, Columbia, MO, 65211, USA, ²Computer Science Department and Christopher S Bond Life Sciences Center, University of Missouri, Columbia, MO, 65211, USA, ³Soybean Genomics and Improvement Laboratory, USDA-ARS, Beltsville, MD, 20705, USA and ⁴MSA Genomics Laboratory, USDA-ARS, Stoneville, MS, 38776, USA

Email: Sandra Thibivilliers - st5y7@mizzou.edu; Trupti Joshi - joshitr@missouri.edu; Kimberly B Campbell - campbelk@ba.ars.usda.gov; Brian Scheffler - bscheffler@msa-stoneville.ars.usda.gov; Dong Xu - xudong@missouri.edu; Bret Cooper - cooperb@ba.ars.usda.gov; Henry T Nguyen - nguyenhenry@missouri.edu; Gary Stacey* - staceyg@missouri.edu

* Corresponding author

**Abstract**

**Background:** *Phaseolus vulgaris* (common bean) is the second most important legume crop in the world after soybean. Consequently, yield losses due to fungal infection, like *Uromyces appendiculatus* (bean rust), have strong consequences. Several resistant genes were identified that confer resistance to bean rust infection. However, the downstream genes and mechanisms involved in bean resistance to infection are poorly characterized.

**Results:** A subtractive bean cDNA library composed of 10,581 unisequences was constructed and enriched in sequences regulated by either bean rust race 41, a virulent strain, or race 49, an avirulent strain on cultivar Early Gallatin carrying the resistance gene *Ur-4*. The construction of this library allowed the identification of 6,202 new bean ESTs, significantly adding to the available sequences for this plant. Regulation of selected bean genes in response to bean rust infection was confirmed by qRT-PCR. Plant gene expression was similar for both race 41 and 49 during the first 48 hours of the infection process but varied significantly at the later time points (72–96 hours after inoculation) mainly due to the presence of the *Avr4* gene in the race 49 leading to a hypersensitive response in the bean plants. A biphasic pattern of gene expression was observed for several genes regulated in response to fungal infection.

**Conclusion:** The enrichment of the public database with over 6,000 bean ESTs significantly adds to the genomic resources available for this important crop plant. The analysis of these genes in response to bean rust infection provides a foundation for further studies of the mechanism of fungal disease resistance. The expression pattern of 90 bean genes upon rust infection shares several features with other legumes infected by biotrophic fungi. This finding suggests that the *P. vulgaris-U. appendiculatus* pathosystem could serve as a model to explore legume-rust interaction.

Published: 27 April 2009

Received: 17 November 2008

Accepted: 27 April 2009
Background

Common bean, *Phaseolus vulgaris*, represents a great source of nutrition for millions of people and is the second most important legume crop, after soybean. It is the target of multiple pests and diseases causing substantial losses. For example, on susceptible bean cultivars, bean rust, caused by *Uromyces appendiculatus*, may cause yield reduction from 18 to 100% with favorable environmental conditions, such as high moisture and temperature between 17 and 27°C [1]. Among the 5 different stages of the bean rust life cycle, basidia, pycnia, aecia, uredinia, and telia, the most devastating on bean is the uredinial stage. The latent period between the germination of an urediniospore and the formation of a sporulating pustule stage can be as short as 7 days. Signs of infection by *Uromyces appendiculatus* include the presence of uredinia or spore-producing pustules on the surface of the leaf. The identification of fungal proteins from quiescent and germinating uredospores enhanced the understanding of the infection process of this fungus [2,3].

Based upon mapping and quantitative trait loci (QTL) analysis, several genes involved in *Colletotrichum lindemuthianum* (Co; anthracnose) resistance and other resistance genes for bean common mosaic virus (BCMV), bean golden yellow mosaic virus (BGYMV), common bacterial blight, and bean rust are clustered [2,3]. The large number of resistance (R) genes for bean rust may correlate with the high pathogen population diversity; with 90 different races identified [4]. The locus *Ur-3* confers resistance to 44 out of the 89 *U. appendiculatus* races present in the USA [5,6]. Besides the *Ur-3* locus, a number of other R genes were identified in bean; such as locus *Ur-4* for race 49, locus *Ur-11* epistatic to *Ur-4* for race 67 or locus *Ur-13* mapped to the linkage group B8 [7,8]. To date, no large scale transcriptomic analysis of bean rust infection has been performed to better understand the mechanism of resistance. All of these *Ur* genes are effective against one specific rust strain, following the gene-for-gene resistance theory. Consequently, gene pyramiding was used to produce cultivars carrying multiple resistance genes [9]. Unfortunately, such resistance may prove to be effective in the field for only a short time due to the adaptation of the fungus to overcome plant defenses [10]. Consequently, unraveling and understanding the mechanisms downstream of these R genes is a key research goal to circumvent the adaptation of the fungus to plant resistance.

We investigated the *Phaseolus vulgaris-Uromyces appendiculatus* pathosystem at a transcriptional level for a better understanding of the plant response to fungal infection. In this study, we developed a subtractive suppressive hybridization (SSH) library made from the common bean cultivar Early Gallatin that exhibits susceptibility to *U. appendiculatus* race 49 (virulent strain) but resistance to *U. appendiculatus* race 49 (avirulent strain). The resistance to *U. appendiculatus* is conferred by the presence of the *Ur-4* gene in this cultivar that leads to a hypersensitive response (HR) in presence of the pathogen race 49 [11]. This cDNA bean library was enriched in expressed sequence tags (ESTs) that are potentially up-regulated by the compatible and incompatible interactions. More than 20,000 clones from the SSH library were sequenced and assembled into contigs. A total of 10,221 *P. vulgaris* sequences and 360 *U. appendiculatus* sequences were added to the NCBI database, significantly increasing the number of ESTs available for common bean. The regulation of 90 genes was confirmed by quantitative real-time polymerase chain reaction (qRT-PCR) revealing 3 main expression patterns and highlighting gene regulation that occurs downstream of R protein activation.

Results and discussion

Identification of unsequences from tissues infected with virulent or avirulent bean rust

Common bean is a diploid (n = 11) with a small genome size estimated at 450 to 650 Mb [12]. So far, the total number of common bean ESTs available is 83,448 (verified on March, 2009). This number was significantly less before the publication of [13] who added ESTs from root nodules, phosphorus deficient roots, developing pods, and leaves, and from leaves and shoots with and without *C. lindemuthianum* inoculation [14]. (The current number also includes the 10,221 ESTs added in this study.). The lack of sufficient *P. vulgaris* sequences precludes the construction of a useful DNA microarray for this plant. Consequently, in order to study the response of bean to *U. appendiculatus* infection, we created a SSH library and sequenced 20,736 clones from 3' and 5' ends. From 41,472 sequences, 8.5% were discarded due to the absence of a cloned sequence or low sequence quality.

During cDNA generation, sequence tags were incorporated prior to pooling cDNAs from different conditions (see Material and Methods for details). The tags identify the treatment and time points used in generating the original mRNA. The distribution of these tags among the library is presented in the Figure 1. Approximately 17% of the sequences lacked a tag after sequencing, while 31% of the sequences had a "race 49" tag and more than 51% had a "race 41" tag. It is important to note that the majority of the ESTs coming from the fungus were tagged "late41", consistent with an effective colonization of the leaf by the virulent fungus (race 41). The lack of tag identification may come from inefficient incorporation of the tag during the library construction or the presence of non-identified nucleotide in the tag sequence making it indiscernible. The various cDNAs in the library could be resolved back to their source tissue by the presence of unique sequence tags. For example, 51% percent of the EST sequences were
derived from bean tissue infected with race 41 since they had the "race 41" tag. This likely reflects the compatible interaction between the race 41 and its host allowing greater fungal penetration. Biotrophic fungi are known to reprogram the host plant cell to support their growth [15] and plant ESTs tagged with race 41 could be involved in this process.

Contig assembly and removal of redundant sequences was performed on 38,592 sequences using TIGR Gene Indices Clustering Tools (TGICL) and CAP3 software. Two thousand seven hundred twenty one sequences showed no similarity with other sequences and were categorized as singletons. These sequences had an average length of 670 bp. Seven thousand eight hundred sixty contigs were assembled from the remaining 35,163 sequences. The average contig length was estimated to be 1 kb. An average contig contains 4.5 sequences (min:2, max:496). Ultimately, 10,581 unisequences were identified and represent genes that are potentially differentially up-regulated during bean infection by virulent or avirulent pathogen isolates.

Among these 10,581 unisequences, 10,221 were annotated as bean genes and 360 were annotated as fungal genes based on best Blast hits to the database These 360 fungal unisequences included 62 singletons and 298 contigs (Table 1).

**Functional annotation**

Sequence analysis revealed that 8,806 ESTs had significant similarity with sequences in public databases such as DFCI or NCBI (using BlastN with an E-value ≤ e-20). Forty-three percent of annotations were based on similarities to sequences in soybean databases and 32.8% were derived from comparisons with common bean (Table 1, see additional file 1: excel file of the 10,581 unisequences). These unisequences were grouped into 18 different functional categories (Figure 2). The most abundant category contained the unknown (31.9%), non-classified (4.7%), and low or no hit (13.7%) groups and represents 50.3% of the entire library. The remaining 49.7% of the sequences were grouped into 14 categories, such as, protein metabolism and catabolism (7.9%), nucleotide and nucleic acid metabolism (8.8%), or stress defense response (2.9%). Taken together, signal transduction regulation and nucleotide and nucleic acid metabolism represent 15.3% of the library. Tian et al. (2007) also found that 14% of their EST library, made from phosphorus starved bean plants, fell into these two categories [16]. Similar observations were made on soybean in response to stresses such as drought, phosphorus starvation or nematode infection [17]. It would seem that under various biotic and abiotic stresses, plants activate several common pathways that alter the expression profile of genes, which allow the plant to respond to the specific environmental condition.

The SSH library was normalized to reduce the redundancy of the most highly expressed genes. However, some genes are very highly expressed and thus remain overrepresented in the normalized library. As expected, the largest contigs (i.e., composed of the most sequences) are involved in basic metabolism processes. The primary metabolism category comprises the largest component (5.9%) of the library based upon the number of unisequences and the proportion of contigs composed of a high number of sequences. For example, CL1contig48 is composed of 496 aligned sequences and not surprisingly, represents ribulose-1,5-bisphosphate carboxylase/oxygenase activase (RuBisco activase). Three of the other 15 largest contigs are also found in the primary metabolism category (Table 2, see additional file 1: excel file of the 10,581 unisequences).

This library was constructed to reveal the plant and fungal genes up-regulated during the rust infection process. Contigs correlated with stress response pathways also have a high number of sequences such as the contig CL1contig105 with 72 sequences encoding an 1-aminoc-
sequences present a perfect match, 762 sequences have 1 mismatch or indel, 658 have 2 mismatches and/or indels, and 1,807 have more than 3 mismatches and/or indels. An average of 1 SNP/indel is putatively identified every 335 bp. However, we were not able to further confirm these SNP/indels due to the lack of the sequence trace files for the bean ESTs present in the NCBI database. This SNP frequency is very similar to that reported previously by Ramirez et al., (2005) who found 1 SNP every 387 bp. Our estimation is based on the comparison of cv Early Gallatin with 3 other cultivars (Bat93, Negro Jamapa 81, and G19883). When this comparison is made between only 2 different cultivars (Early Gallatin and G19883) the SNP frequency in the coding sequences decreases to 1 SNP every 570 bp. For comparison, the SNP frequency in the soybean coding sequence was estimated at 1 SNP/490 bp in exons and 1 SNP/375 bp in introns [18]. The genes identified by EST sequencing represent candidates involved in the plant host's ability to withstand rust infection. Therefore, genetic mapping of these gene candidates is a means to correlate their position with known QTL involved in disease resistance.

The 360 fungal sequences represent 3.4% of the library. Two studies in rice showed that the harvesting time (i.e., fungal biomass in the infected leaf is low at the earliest time points) and the stringency of selection (i.e., choice of the appropriate E-value for the blast) are very important to accurately sample the abundance of fungal EST in infected leaf tissue [19,20]. In this study, the selected E-value was e-20, greatly reducing the risk of false positive clones. Tissue was sampled after 5 days of infection allowing the multiplication of the fungi in the leaf tissue. At 5DAL, the haustoria are already mature and are probably redirecting the nutrient uptake from the plant based on their genes expression pattern [21].

These genes were mainly annotated predominantly by comparison to ESTs from germinating uredospires of *U. appendiculatus* [22,23] (Table 3, see additional file 1: excel file of the 10,581 unisequences). Two hundred seventy two fungal sequences, representing 74.4% of the total, were considered identical to ESTs already present in the NCBI database while 88 sequences are new and unique as identified by less than 98% identity over at least 100 bp. Interestingly, among the 88 fungal ESTs that showed no similarity with ESTs from *U. appendiculatus* germinating uredospires, 19 showed similarity with *Uromyces viciae* haustorium-specific cDNAs and may be specific to successful infections. These remaining sequences represent candidates for fungal genes more directly involved in the infection mechanism. The library was made from tissues infected with a virulent and avirulent rust strain to allow for the identification of genes involved in both pathogen-host compatibility and resistance. Beside, the high simi-

| EST Distribution Analysis | Unisequence | singleton | contig = 2 | contig>2 |
|---------------------------|-------------|-----------|------------|----------|
| vector                    | 13          | 2         | 7          | 4        |
| No hit                    | 118         | 49        | 66         | 3        |
| Low hit (>e-20)           | 1279        | 514       | 663        | 102      |
| Fungi                     | 360         | 62        | 210        | 88       |
| Barley                    | 1           | 1         | 0          | 0        |
| Ice plant                 | 1           | 0         | 1          | 0        |
| Petunia                   | 1           | 1         | 0          | 0        |
| *S. pombe*                | 1           | 0         | 1          | 0        |
| Sugarcane                 | 1           | 1         | 0          | 0        |
| Sunflower                 | 1           | 0         | 1          | 0        |
| Beet                      | 2           | 1         | 0          | 1        |
| *N. benthamiana*          | 2           | 1         | 1          | 0        |
| Oilseed rape              | 2           | 0         | 0          | 2        |
| Tobacco                   | 2           | 1         | 1          | 0        |
| *Pinus*                   | 3           | 3         | 0          | 0        |
| Potato                    | 3           | 2         | 1          | 0        |
| Lettuce                   | 4           | 2         | 2          | 0        |
| *Wheat*                   | 5           | 2         | 0          | 3        |
| Aquilegia                 | 6           | 3         | 2          | 1        |
| *Rice*                    | 6           | 5         | 0          | 1        |
| Tomato                    | 7           | 0         | 5          | 2        |
| *A. thaliana*             | 9           | 4         | 3          | 2        |
| Poplar                    | 15          | 6         | 5          | 4        |
| Grape                     | 18          | 6         | 9          | 3        |
| Cotton                    | 23          | 8         | 13         | 2        |
| Lotus                     | 145         | 51        | 52         | 42       |
| *M. truncatula*           | 500         | 190       | 184        | 126      |
| Common bean               | 3473        | 622       | 1249       | 1602     |
| Soybean                   | 4580        | 1184      | 1844       | 1552     |
| Total                     | 10581       | 2721      | 4320       | 3540     |

"Unisequence", "singleton", "contig = 2", and "contig>2" columns represent the number of total unisequences, those found once in the SSH library, the contigs made up of only 2 sequences and the contigs composed of more than 2 sequences, respectively, having a hit with an organism listed in column one.
larity of these 19 sequences with haustoria-specific ESTs makes them likely candidates to encode potential effectors or avirulence proteins.

The largest contig has sequence similarity to a putative beta-galactosidase (an enzyme involved in the degradation of the cell-wall) based on a match to a cDNA from germinating *P. pachyrhizi* uredospores.

**Comparative analysis with the Phaseolus vulgaris BAC-end sequences**

Recently, the University of Purdue released the first bean FingerPrinted Contigs (FPC) physical map that contains 7,567 contigs or singletons and is anchored with 240 genetic markers [http://phaseolus.genomics.purdue.edu/](http://phaseolus.genomics.purdue.edu/). The Bacterial artificial chromosome-ends (BAC) of the clones were sequenced and provide a powerful tool for integrated genomic and genetic analysis. This recent release of the *Phaseolus* physical map [http://phaseolus.genomics.purdue.edu/](http://phaseolus.genomics.purdue.edu/) allowed linkage of some of the 10,221 unisequences to the physical map by comparison to BAC-end sequences. The number of ESTs matching a BAC-end sequence was assessed using the minimal criteria of 98% sequence identity spanning 100 bp. As a result, 1,704 unisequences, more than 15% of the total library, could be linked to the physical map (see additional file 2: excel file of the 1,704 unisequences having a hit with a BAC-end sequence). Fourteen of these sequences located to physical contig 1 that is composed of 999 BACs. The genetic mapping of these ESTs might be facilitated by the presence of genetic markers anchoring BACs within the various contigs.

This physical map was made from common bean cultivar G19833. The number of putative SNP between the BAC-end sequences and the ESTs (common bean cultivar Early gallatin) was identified among the 1,704 matching ESTs. Six hundred seventy ESTs showed a perfect match with a BAC-end sequence, 414 ESTs exhibited only 1 mismatch, 258 ESTs contained 2 mismatches and 362 ESTs had 3 or more mismatches. This represents an average of 1 SNP/indel every 570 bp.

**Identification of bean reference genes**

Among the 10,221 unisequences, we sought to confirm the expression of 90 ESTs using qRT-PCR. To normalize gene expression based on qRT-PCR, the identification of constitutively expressed bean reference genes is required. The use of house keeping genes as reference genes for gene expression normalization can induce some error in the analysis of the data without confirmation of their constitutive expression especially when using qRT-PCR [24,25]. Consequently, three bean genes, *TC197*, *TC127*, and *TC185* (encoding a guanine nucleotide-binding protein...
beta subunit-like protein, ubiquitin, and tubulin beta chain respectively) were selected based on their housekeeping function and/or their presence in different bean cDNA libraries [13,14]. Additionally, homologs of soybean genes cons6, cons7, and cons15 (encoding for a F-box protein family, a metalloprotease, and a peptidase S16 respectively), were chosen since they were recently shown to be expressed constitutively in soybean [26].

Preliminary analysis of these putative constitutive genes by qRT-PCR performed on leaf, stem, and pod cDNA led to the elimination of TC197, cons15 and TC1985 due to the variability of their expression levels (data not shown). The stability of the expression level of the 3 remaining genes, TC127, cons6 and cons7 was evaluated by qRT-PCR on cDNAs from bean uninoculated or infected with bean rust race 41 or 49 at 6, 12, 24, 48, 72, and 96 hours after inoculation (HAI). After analysis of their expression stability using geNorm software [27], cons7 was the most stably expressed in our experimental conditions (Figure 3). For this reason, cons7 was selected for normalization of the expression data. It is interesting to note that cons7 was also among the most stably expressed constitutive genes in soybean [26] and, therefore, could be a candidate to use for expression normalization in other legumes.

**Transcriptional analysis of selected ESTs during the bean rust infection process**

In order to compare expression of genes responding to *U. appendiculatus* race 41 to those responding to race 49, during bean infection and colonization, the expression level of six, selected fungal genes was analyzed using qRT-PCR (Figure 4). During the first 24 hours of the infection, the six genes were expressed at comparable levels. However, by 48 HAI, the expression of all six genes was significantly higher in tissues infected with the virulent race 41 isolate. This result likely reflects the nature of the compatible, virulent interaction as compared to inhibition of race 49 infection by the host defenses. Consistent with this, all six genes used in this analysis came from the ESTs possessing the tag of the "late41" library. The EST CL3018Contig1, encoding for a plant-induced rust protein 1, exhibits significant similarity with NMT1 (no messenger in thiamine), which is involved in the biosynthesis of the pyrimidine moiety of thiamine (vitamin B1). This gene was strongly expressed only in tissue infected with the virulent fungus race 41. Similar observations were made previously using bean plants infected with *Uromyces fabae* [28]. These data also suggest that the haustoria may not only be the site of nutrient uptake from the plant [29] but also the site of metabolite biosynthesis with specific haustorial genes involved in vitamin biosynthesis [e.g., NMT1] [28].

Ninety bean unisequences were selected (based on their putative function and tag) and their regulation was confirmed by qRT-PCR using RNA obtained from three independent biological replicates. Unisequences coming from the ESTs in the "race 49" tagged libraries were desirable due to their potential involvement in a resistance pathway. The regulation of these genes was evaluated by qRT-PCR using RNA from un inoculated leaf tissues or those...
inoculated with either *U. appendiculatus* uredospores of race 41 or race 49 isolates at the time points 0, 6, 12, 24, 48, 72, or 96 HAI. The data obtained were used to compare the ratio of gene expression in tissues infected with race 41 or race 49 to that in uninoculated bean leaves. The data also allowed a direct comparison of gene expression induced by either race 41 or race 49. The first two comparisons highlight regulation in the infected plants by the rust fungi, while the third comparison highlights gene expression differences between the two types of infection, resistant and susceptible. The 90 genes showed significant expression differences in at least one of the 3 comparisons (p-value < 0.05, cut-off < -1 or > 1 or p-value < 0.1, cut-off < -0.58 or > 0.58 in log base 2).

The transcriptional response was profiled in relation to the time after inoculation (Figure 5). For example, 39 and 41 genes showed differential regulation within 6 and 12 HAI, respectively, in tissue inoculated with race 41. At these same time points 40 and 24 genes, respectively, were differentially regulated in tissues infected with race 49. At the latest time points, 72 and 96 HAI, 16 and 19 genes, respectively, for race 41 and 6 and 14 genes, respectively, for race 49 were differentially regulated. It is interesting to note that the regulation occurring at the early time points appeared to be independent of the fungal race used for inoculation. At the early time points (i.e., first 48 hours), only 16 genes (36% of those tested) showed a difference in expression in tissues inoculated with the two fungal races. However, at the later time points, this number increased to 34 genes with 18 (36%) and 16 (32%) at 72 and 96 HAI, respectively. These results suggested that during the beginning of the infection most of the bean gene regulation is independent of the fungal race, but differences due to fungal race occur as the infection progresses. It is possible that fungal-Pathogen Associated Molecular Pattern (PAMP) elicitors (e.g., chitin) induce the same response from the plant at the beginning of the infection. Subsequently, the Avr4 protein in the race 49 is recognized after a couple of days leading to the induction of defense-related genes. However, in bean infected by race 41, no plant defense is activated and gene expression may reflect the reprogramming of the plant host by the fungus especially at the haustorial site.

A key finding of the van de Mortel *et al.* (2007) study on *Glycine max* – *Phakopsora pachyrhizi* was that most genes were regulated early during infection (the first 24 hours) and at the latest time points tested (72–120 hours). However, at the intermediate time points (24 to 72 hours), few genes were regulated; this phenomenon was called a “dip” by van de Mortel *et al.* (2007). This same expression profile was also observed in bean upon rust infection by both races. At 24 and 48 HAI few genes were regulated in comparison to 6–12–72 and 96 HAI. In common bean, those genes regulated at 6 HAI were very different from those

---

### Table 3: List of the 15 most common fungal rust contigs containing the highest number of sequences

| Total ESTs in contig | Contig annotation | genus | gene annotation *
|---------------------|-------------------|-------|-----------------|
| 18                  | CL1Contig467      | *Phakopsora pachyrhizi* | cDNA from germinating uredospores SSH-library similar to beta-galactosidase |
| 14                  | CL229Contig1      | *Uromyces viciae*       | cDNA from hyphae from germinating uredospore similar to translation elongation factor |
| 14                  | CL201Contig1      | *Uromyces appendiculatus* | cDNA from hyphae from germinating uredospore similar to von Willebrand factor |
| 14                  | CL1Contig310      | *Uromyces appendiculatus* | cDNA from hyphae from germinating uredospore similar to von Willebrand factor |
| 10                  | CL124Contig1      | *Uromyces appendiculatus* | cDNA from hyphae from germinating uredospore similar to unknown |
| 10                  | CL116Contig1      | *Uromyces appendiculatus* | cDNA from hyphae from germinating uredospore similar to unknown |
| 10                  | CL492Contig1      | *Uromyces appendiculatus* | cDNA from hyphae from germinating uredospore similar to unknown |
| 8                   | CL633Contig1      | *Uromyces viciae*       | haustorium-specific cDNA similar to nucleotide excision repair protein yeast rad23 |
| 8                   | CL662Contig1      | *Uromyces viciae*       | haustorium-specific cDNA similar to 6-phosphogluconate dehydrogenase |
| 8                   | CL16Contig2       | *Uromyces appendiculatus* | cDNA from hyphae from germinating uredospore similar to unknown |
| 8                   | CL766Contig1      | *Puccinia graminis f. sp. tritici* | SSH-library of Puccinia graminis infected wheat leaves similar to 60s ribosomal protein L5 gene |
| 8                   | CL772Contig1      | *Uromyces viciae*       | haustorium-specific cDNA similar to voltage-dependent ion-selective channel |
| 8                   | CL582Contig1      | *Puccinia graminis f. sp. tritici* | SSH-library of Puccinia graminis infected wheat leaves similar to glutathione S-transferase |
| 8                   | CL787Contig1      | *Uromyces appendiculatus* | cDNA from hyphae from germinating uredospore similar to cysteine-rich secretory protein (CRISP/SCP/TPX1) |

* Based on the BlastX hits (E value > e^-20) with fungal ESTs
expression is subsequently maintained only in the plants infected by race 49 at the later time points (up to 96 HAI). Genes in this group include those annotated as plant defense (35% of this group) containing PR1, wound induced protein 2 (WIN2) genes, cell-wall related (i.e., a cell-wall invertase gene), and signal transduction regulation category with a G-box binding protein PG2 or sensory transduction histidine kinase genes. These genes are likely involved in the defense pathways induced by a fungal-PAMP since they have the same expression pattern during the first 24 hours of infection independent of the fungal race used. For example, the wound-induced protein WIN2 protein has anti-fungal activity [32] and possesses a domain that can bind a well known PAMP, chitin [33]. The formation of haustoria by the fungus in the plant can occur within hours of infection [34]. After successful colonization of the bean cell, rust race 41 likely secretes effector proteins that can suppress the plant defense pathway induced by PAMPs. The initial induction of genes such as WIN2 by race 41 and their subsequent reduction in expression may be associated with this suppression of defense by the virulent pathogen only. The second group, group B, is composed of 16 genes that were induced at the beginning of the infection but were slightly down-regulated at the later time points independent of the fungal race. This group is rich in genes categorized as plant defense representing 56% of this group. The third group, group C (C1 and C2), is composed of 5 genes that appeared to be repressed by inoculation. In group C1, the genes were repressed during the first 12 hours by both races but this repression was only maintained at later time points (i.e., 72 and 96 HAI) in tissues infected with race 49. In contrast to group C2, the apparent repression of genes occurred only after 72 HAI with both races. The fourth group, group D, consists of 35 genes that were repressed in the first 12 hours by both races and subsequently expressed at levels comparable to the uninoculated tissue. The genes repressed specifically at the early time points could also be involved in the basal defense response. This pool is composed of ESTs known to be involved in plant defense pathways [e.g., chitinase class 1 [35], an auxin response factor 4 and an auxin conjugate hydrolase [36], and a MLO-like protein 8 [37]]. Finally, the fifth group represents 18 genes that gave no discernable pattern of expression.

A more detailed analysis was performed on the expression ratio of transcripts in bean leaves inoculated with the fungus race 41 or race 49 versus uninoculated bean leaves. These analyses are presented in a hierarchical cluster based on Euclidian distance (Figure 6, see additional file 3: excel file of the ratio of the expression level of the 90 regulated genes for all conditions). This cluster can be divided into five main groups. The first, group A (A1 and A2), is composed of 17 genes up-regulated by both fungal races in the first 24 hours of the infection but enhanced

Figure 3
Ranking of bean genes based on their expression stability measured by qRT-PCR. The expression levels of three putative constitutive genes (TC127, cons6, and cons7) was measured during infection by both fungal race 41 and 49 in order to identify the best reference gene for qRT-PCR normalization. Genes with the most stable expression during the conditions tested are on the right of the diagram, the less stably expressed being on the left. Figure generated by geNorm software.
The same observation was made also by Lee et al. (2008) at the protein level. Lee et al. (2008) proposed a new model for plant disease resistance where R-gene mediated resistance is integrated into the basal immunity system of the plant and functions primarily to restore the innate immunity response that is actively suppressed by virulent pathogens[38]. Similar patterns of expression, independent of the pathogen virulence, were observed in Arabidopsis [39] and barley [40]. Another category of genes (i.e. cell-wall invertase or amino acid transporter-like protein 1) involved in the plant defense system are likely involved in the HR and were regulated only at the later time points in plants infected by the race 49 fungus. The expression of these genes may be the result of recognition of AvrUR-4 by the Ur-4 resistance protein and lead to the presence of HR ten days after infection with this isolates.

**Conclusion**

In summary, we identified 10,581 P. vulgaris unisequences and confirmed the regulation of 90 plant genes by rust infection in common bean. These data have added significantly to the genomic resources available for common bean, while also providing insight into how this plant responds to fungal infection. As part of this study, we identified constitutively expressed bean genes that can be used for normalization in gene expression studies. The data also suggest that a biphasic gene expression pattern may be a common feature in plants infected by biotrophic fungi.

**Methods**

**Plant material**

Bean tissues were produced at the USDA-ARS facility (Beltsville, MD). P. vulgaris cv. Early Gallatin plants were inoculated with either U. appendiculatus race 41 (virulent strain) or race 49 (avirulent strain) uredospores. The primary leaves of 10 day old plants were inoculated on the top and bottom. Spores (2 × 10^5 spores/ml) were mixed in water and then sprayed on leaves with an aerosol canister. The plants were placed in a dew chamber in the dark at 20°C for 12 hours and then moved to a growth room for 12 hours.
using genetic research facility (USDA-ASR, Stoneville, MS). The 3730xl DNA sequencer (AME Bioscience) at the catfish The 20,736 cDNA clones were sequenced using an ABI method [41]. The cDNA from bean infected with described by Bonaldo composed of more than 20,000 ESTs and was prepared as Carver Biotechnology Center (Urbana, IL). The library is The normalized SSH library was generated at the Roy J. lesions when inoculated with race 41 or race 49. The enrichment of cDNA derived from uninoculated leaves and germinated spores. The library was subsequently sub-divided in 4 pools against cDNA from the 4 described pools against cDNA derived from uninoculated leaves and germinated spores. The library was subsequently sub-divided in 4 parts based on sequence tags added during library construction. 20,736 cDNAs were cloned in pGem-T (Promega) for sequencing.

**Sequencing and data processing**
The 20,736 cDNA clones were sequenced using an ABI 3730xl DNA sequencer (AME Bioscience) at the catfish genetic research facility (USDA-ASR, Stoneville, MS). The conversion of the electropherogram into base and quality files was performed using Phred [42]. The EST sequences were first cleaned of polyA, polyT, and low complexity sequence using SeqClean from TIGR http://compbio.dfci.harvard.edu/tgi/software/. Contig assembly was done using the TIGR Gene Indices Clustering Tools (TGICL) http://compbio.dfci.harvard.edu/tgi/software/ after removing vector and tag sequences. It uses a slightly modified version of NCBI’s megablast, and the resulting clusters are then assembled using the CAP3 assembly program (Huang and Madan. 1999). Annotations for the sequences were obtained by Blast against the TIGR plant and fungal sequence databases and Uniprot database. The ESTs were submitted to NCBI Genbank dbEST under the accession numbers FE674093 to FE712011.

**RNA extraction**
RNA extraction and cDNA synthesis from leaf tissues of common bean cv Early Gallatin infected with either *U. appendiculatus* race 41 (virulent strain) or race 49 (avirulent strain) and from soybean tissues infected by *P. pachyrhizi*, were performed as described by Libault et al., 2008. Briefly, RNAs were extracted from ground frozen tissues using TRIzol@reagent (Invitrogen, Carlsbad, Calif.) and purified by two phenol/chloroform extractions. The RNAs were treated with TURBO DNA-free enzyme (Ambion) to remove all DNA contaminants. cDNA synthesis was prepared from 5 μg of RNA using the MMLV reverse transcriptase (Promega, Madison, WI).

**Quantitative PCR Primer Design**
The qRT-PCR primers were designed with primer3 software http://frodo.wi.mit.edu/primer3/input.htm using the following criteria, Tm of 60°C, PCR amplicon length from 80 to 125 bp, primer sequence length from 19 to 23 nucleotides with guanine-cytosine contents from 40% to 60% (see additional file 4: excel file of the qRT-PCR primers).

**qRT-PCR reaction conditions and data analysis**
The qRT-PCR on bean leaf tissues were performed in a 384-well plate format (7900 HT Sequence detection System; Applied Biosystems, Foster City, CA). The qRT-PCR of soybean leaf, pod, and stem tissues was performed with a 96-well plate qRT-PCR machine (7500 Real-Time PCR System; Applied Biosystems, Foster City, CA). Data analysis was performed as described by Libault et al. (2008) with modifications [43]. The data collection was performed during 40 cycles for bean but 45 cycles for soybean with an Rn threshold set at 0.2 for Ct value determination. The ratios of the expression level were transformed into a Log 2 base for clustering in Gene traffic software using a hierarchical clustering algorithm. A t-test was used to assess the statistical differences of the mean of the ratio for each sample at each time point.
Figure 6
Transcription profile of the bean genes during rust infection. Hierarchical clustering of the expression ratio of the ESTs differentially expressed in the inoculated bean by race 41 or 49 in comparison to uninoculated bean tissue. qRT-PCR was performed on 3 independent biological replicates. Analyses were performed on the data set in a Log 2 form based on Euclidean distance.
The identification of a reference gene for qRT-PCR normalization was made using geNorm software [27]. This software calculates the mean pairwise variation (based on geometrical mean) for each gene and compares these values among these genes. A high mean pairwise variation is found for gene with low expression stability.

**Authors' contributions**

ST contributed to the sequences production and analysis, performed the expression analysis and drafted the manuscript; TJ performed the bioinformatic analysis. KBC produced the plant material and contributed to the sequences production; BS sequenced the clones; DX supervised the bioinformatic part of the project and worked over the draft version of the manuscript; BC and HTN supervised the project and worked over the draft version of the manuscript; GS conceived and supervised the project and worked over the draft version of the manuscript. All authors read and approved the final manuscript.

**Acknowledgements**

We are grateful to Marc Libault, William Franck, and James English for their critical reading of this manuscript. This research was supported by DOE grant DE-PS02-07ER07-02.

**References**

1. Mnmbaga MT, Steadman JR, Stavely JR: The use of host resistance in disease management of rust in common bean. *Integrated Pest Management Reviews* 1996, 1:191-200.
2. Rodríguez-Suárez C, Mendez-Vigo B, Pineda A, Ferreira JJ, Giraldes R: A genetic linkage map of Phaseolus vulgaris L. and localization of genes for specific resistance to six races of anthracnose (Colletotrichum lindemuthianum). *TAG Theoretical and applied genetics* 2007, 114(4):713-722.
3. Mendoza A, Hernández H, Hernández S: Identification of Co-1 Anthracnose Resistance and Linked Molecular Markers in Common Bean Line A193. *Plant Disease* 2001, 85(3):252-255.
4. Araya CM, Alleyne AT, Steadman JR: Pheno-typic and genotypic characterization of Uromyces appendiculatus from Phaseolus vulgaris in the Americas. *Plant Disease* 2004, 88(8):830-836.
5. Sayler RJ, Ewing JD, McClean PE: Monogenic and epistatic resistance to bean rust infection in common bean. *Physiological and Molecular Plant Pathology* 1995, 47(3):173-184.
6. Stavely JR: Development of comprehensive rust resistant bean gene pools (Abstr.). *Physiological and Molecular Plant Pathology* 1998, 53(supp):585.
7. Kalavacharia V, Stavely JR, Myers JR, McClean PE: Crg, a gene required for Ur-3-mediated rust resistance in common bean, maps to a resistance gene analog cluster. *Mol Plant Microbe Interact* 2000, 13(11):1237-1242.
8. Mienie CM, Liebenberg MM, Pretorius ZA, Mildas PN: SCAR markers linked to the common bean rust resistance gene Ur-13. *TAG Theoretical and applied genetics* 2005, 111(5):972-979.
9. Pastor-Corrales MA: Sources, genes for resistance and pedigrees of 52 rust and mosaic resistant dry bean germplasm lines released by the USDA Beltsville bean project in collaboration with Michigan, Nebraska and North Dakota Agricultural Experiment Stations. *Ann Rev of the Bean Improv Coop* 2003, 46:235-241.
10. Johnson R: A critical analysis of durable resistance. *Annu Rev Phytopathol* 1984, 22:309-330.
11. Pastor-Corrales MA, Stavely JR: Using specific races of the common bean rust pathogen to detect resistance genes in Phaseolus vulgaris. *Bean Improvement Cooperative. Annual Report* 2002, 45:78-79.
12. Bennett MD, Leitch IJ: Nuclear DNA amounts in angiosperms. *Ann Bot (London)* 1995, 76:113-116.
13. Ramirez M, Graham MA, Blanco-Lopez L, Silvente S, Medrano-Soto A, Blair PW, Hernandez G, Vanez CP, Lara M: Sequencing and analysis of common bean ESTS. Building a foundation for functional genomics. *Plant physiology* 2005, 137(4):1211-1227.
14. Melotto M, Monteiro-Vitorello CB, Bruschi AG, Camargo LE: Comparative bioinformatic analysis of genes expressed in common bean (Phaseolus vulgaris L.) seedlings. *Genome/National Research Council Canada = Genome/Conseil national de recherches Canada* 2005, 48(supp):562-570.
15. Chou HM, Bundock N, Rolfe SA, Scholes JD: Infection of Arabidopsis thaliana leaves with Albugo candida (white blister rust) causes a reprogramming of host metabolism. *Molecular Plant Pathology* 2000, 1(2):99-113.
16. Tian J, Venkatachalam P, Liao H, Yan X, Raghothama K: Molecular cloning and characterization of phosphorus starvation responsive genes in common bean (Phaseolus vulgaris L.). *Planta* 2007, 227(1):151-165.
17. Umezawa T, Sakurai T, Tokoki Y, Toyoda A, Seki M, Ishiwata A, Akiyama K, Kurotani A, Yoshida T, Mochida K, et al.: Sequencing and analysis of approximately 40,000 soybean cDNA clones from a full-length-enriched cDNA library. *DNA Res* 2008, 15(6):333-346.
18. Choi YJ, Hyen DL, Matukumalli LK, Song Q, Chaky M, Quigley CV, Chase K, Lark KG, Reiter RS, Yoon MS, et al.: A soybean transcript map: gene distribution, haplotype and single-nucleotide polymorphism analysis. *Genetics* 2007, 176(1):685-696.
19. Jantassiyarar C, Gowda M, Haller K, Hatfield J, Lu G, Stahlberg E, Zhou B, Li H, Kim H, Yu Y, et al.: Large-scale identification of expressed sequence tags involved in rice and rice blast fungus interaction. *Plant physiology* 2005, 138(1):105-115.
20. Kim S, Ahn IP, Lee YH: Analysis of genes expressed during rice-Magnaporthe grisea interactions. *Mol Plant Microbe Interact* 2001, 14(11):1340-1346.
21. Puthoff DP, Neelam A, Ehrenfried ML, Scheffler BE, Ballard L, Song Q, Campbell KB, Cooper B, Tucker ML: Analysis of Expressed Sequence Tags from Uromyces appendiculatus Hyphae and Haustoria and Their Comparison to Sequences from Other Rust Fungi. Phytopathology 2008, 98(10):1126-1135.

22. Cooper B, Garrett WM, Campbell KB: Shotgun identification of proteins from uredospores of the bean rust Uromyces appendiculatus. Proteomics 2006, 6(8):2477-2494.

23. Cooper B, Neelam A, Campbell KB, Lee J, Liu G, Garrett WM, Scheffler B, Tucker ML: Protein accumulation in the germinating Uromyces appendiculatus uredospore. Mol Plant Microbe Interact 2007, 20(7):857-866.

24. Czechowski T, Stitt M, Altmann T, Udvardi MK, Scheible WR: Genome-wide identification and testing of superior reference genes for transcript normalization in Arabidopsis. Plant physiology 2005, 135(1):5-17.

25. Radonic A, Thulke S, Bae HG, Muller MA, Siegert W, Nitsche A: Reference gene selection for quantitative real-time, PCR analysis in virus infected cells: SARS corona virus, Yellow fever virus, Human Herpesvirus-6, Camelpox virus and Cytomegalovirus infections. Virology journal 2005, 2:7.

26. Libault M, Thibivilliers S, Radwan O, Clough SJ, Stacey G: Identification of four soybean reference genes for gene expression normalization. The plant Genome 2008, 1:44-54.

27. Vandesompele J, De Preter K, Pattyn F, Poppe B, Van Roy N, De Paepe A, Speleman F: Accurate normalization of real-time quantitative RT-PCR data by geometric averaging of multiple internal control genes. Genome biology 2002, 3(7):RESEARCH0034.

28. Sohn J, Voegele RT, Mendgen K, Hahn M: High level activation of vitamin B1 biosynthesis genes in haustoria of the rust fungus Uromyces fabae. Mol Plant Microbe Interact 2000, 13(6):629-636.

29. Staples RC: Nutrients of a rust fungus: the role of haustoria. Trends in Plant Science 2001, 6:496-498.

30. Gregersen PL, Thordal-Christensen H, Forster H, Collinge DB: Differential gene transcript accumulation in barley leaf epidermis and mesophyll in response to attack by Blumeria graminis f.sp. hordei (syn. Erysiphe graminis f.sp. hordei). Physiological and Molecular Plant Pathology 1997, 51:85-97.

31. Moy P, Qutob D, Chapman BP, Atkinson I, Gijzen M: Patterns of gene expression upon infection of soybean plants by Phytophthora sojae. Mol Plant Microbe Interact 2004, 17(10):1051-1062.

32. Stanford A, Bevan M, Northcote D: Differential expression within a family of novel wound-induced genes in potato. Mol Gen Genet 1989, 215(2):200-208.

33. Ponstein AS, Bres-Vloemans SA, Sela-Buurlage MB, Elzen PJ van den, Melchers LS, Cornelissen BJ: A novel pathogen- and wound-inducible tobacco (Nicotiana tabacum) protein with antifungal activity. Plant physiology 1994, 104(1):109-118.

34. Mendgen K, Wirsig SG, Lux A, Hoffmann J, Boland W: Volatiles mediate the development of plant pathogenic rust fungi. Pflanze 2006, 224(6):1353-1361.

35. Ji C, Kuc J: Purification and characterization of an acidic beta-1,3-glucanase from cucumber and its relationship to systemic disease resistance induced by Colletotrichum lagenarium and tobacco necrosis virus. Mol Plant Microbe Interact 1995, 8(6):899-905.

36. Navarro L, Dunoyer P, Jay F, Arnold B, Dharmasiri N, Estelle M, Voinnet O, Jones JD: A plant miRNA contributes to antibacterial resistance by repressing axin signaling. Science (New York, NY) 2006, 312(2772):436-439.

37. Stein M, Somerville SC: MLO, a novel modulator of plant defenses and cell death, binds calmodulin. Trends Plant Sci 2002, 7(9):377-380.

38. Lee J, Feng J, Campbell KB, Scheffler BE, Garrett WM, Thibivilliers S, Stacey G, Naiman DQ, Tucker ML, Pastor-Corrales MA, Cooper B: Quantitative proteomic analysis of bean plants infected by a virulent and avirulent obligate rust fungus. Mol Cell Proteomics 2009, 8(1):19-31.

39. Feyes BJ, Wiermer M, Bhat RA, Moisan LJ, Medina-Escobar N, Neu C, Cabral A, Parker JE: Arabidopsis SENESCENCE-ASSOCIATED GENE101 stabilizes and signals within an ENHANCED DISEASE SUSCEPTIBILITY1 complex in plant innate immunity. The Plant cell 2005, 17(9):2601-2613.

40. Shen QH, Sajio Y, Mauch S, Biskup C, Bieri S, Keller B, Seki H, Ulker B, Somssich IE, Schulze-Lefert P: Nuclear activity of MLA immune receptors links isolate-specific and basal disease-resistance responses. Science (New York, NY) 2007, 315(5815):1098-1103.

41. Bonaldo MF, Lennon G, Soares MB: Normalization and subtraction: two approaches to facilitate gene discovery. Genome research 1998, 8(3):175-185.

42. Ewing B, Hillier L, Wendl MC, Green P: Base-calling of automated sequencer traces using phred. I. Accuracy assessment. Genome research 1996, 6(9):791-806.

43. Libault M, Wan J, Czechowski T, Udvardi MK, Stacey G: Identification of 118 Arabidopsis transcription factor and 30 ubiquitin-ligase genes responding to chitin, a plant-defense elicitor. Mol Plant Microbe Interact 2007, 20(8):900-911.