

# Bioinformatic comparative analysis of lovastatin gene cluster in endophytic fungi and a soil fungus, *Aspergillus terreus*

## Abstract

Genes responsible for lovastatin production in certain fungi are often clustered (64kb DNA). In the present study, we report the comparative analysis of the genome of sixty one endophytic fungi and a soil fungus for the presence of lovastatin gene sequences. Nucleotide sequences were obtained from NCBI and aligned with lovastatin gene cluster sequence (AF141924.1 & AF141925.1). DNA sequences that are identical to lovastatin gene were identified in *Aspergillus terreus*, (AH007774.1), a soil isolate, whereas all the other sixty one endophytic fungi including a species of *Aspergillus terreus* showed very limited or no homology with the lovastatin gene cluster.

**Keywords:** lovastatin, gene cluster, alignment, endophytic fungi, soil fungus, *aspergillus terreus*

Volume 1 Issue 4 - 2014

**Bhargavi SD, Praveen VK, Savitha J**

Department of Microbiology and Biotechnology, Bangalore University, India

**Correspondence:** Savitha J, Department of Microbiology and Biotechnology, Bangalore University, Bangalore-560056, Karnataka, India, Tel 08022961461, Email drsvtj@yahoo.co.in

**Received:** August 26, 2014 | **Published:** September 22, 2014

**Abbreviations:** ORF, open reading frames; BLAST, basic local alignment search tool; EMBL, european molecular biology laboratory

## Introduction

Fungi belonging to different genera produce lovastatin, a secondary metabolite that competitively inhibits the enzyme Hydroxyl Methyl Glutaryl Coenzyme A (HMG CoA) reductase, which is involved in the biosynthesis of cholesterol. Lovastatin is an effective drug for treatment of hyperlipidemia, a condition in which there is a considerable elevated level of lipids in blood.<sup>1</sup> Owing to its biomedical importance, research on lovastatin and lovastatin producing microorganisms has intensified in recent years. Specifically, terrestrial filamentous fungi and aquatic filamentous fungi have been explored extensively but reports on fungi inhabiting unusual environment are meagre.

Genes responsible for secondary metabolism in fungi are often linked or clustered. In recent years, there has been an exponential increase in the number of reports on the whole genome sequence of fungi in general, and pharmaceutically important fungi, in particular. Comparison of these sequences with the gene sequence of interest has proved to be an excellent resource and promising approach for recombinant DNA technology.<sup>2</sup>

Studies on structural analysis of the lovastatin gene in *Aspergillus terreus* revealed that there were 18 open reading frames (ORF) within the cloned 64kb DNA; 13 of which had been functionally determined by basic local alignment search tool (BLAST). The gene cluster consists of the lovastatin nonaketide synthase gene (lovB), lovastatin diketide synthase gene (lovF), enoyl reductase gene (lovC), transesterase gene (lovD), HMG-CoA reductase gene (ORF8), regulatory genes (lovE and ORF13) which encodes a GAL4-like transcriptional factor and cytochrome P450 monooxygenase gene.<sup>3</sup>

In our previous study, we screened endophytic fungi isolated from medicinal plants as well as terrestrial fungi isolated from soil for their ability to produce lovastatin using wheat bran as substrate. The isolate that produced copious amounts of lovastatin was identified as *Aspergillus terreus* (KM017963). None of the 54 endophytic fungi

isolated from medicinal plants produced any detectable levels of lovastatin.<sup>4</sup> This present study was initiated to determine if the lack of lovastatin production by endophytic fungi is due to absence of genetic elements associated with lovastatin biosynthesis or due to other physiological factors.

## Materials and methods

A total of sixty two filamentous fungi including two strains of *Aspergillus terreus* were considered for the study. Sixty one out of sixty two isolates were endophytic fungi and one was a soil isolate (*Aspergillus terreus*). Whole genome nucleotide sequences of all the fungi and lovastatin gene cluster sequence (AF141924.1 & AF141925.1) used in the present study were obtained from NCBI genbank database in the FASTA format. Both AF141924.1 & AF141925.1 represents the gene cluster of lovastatin. Pair wise alignment tool of European Molecular Biology Laboratory (EMBL-EBI) was used to compare lovastatin gene cluster sequence with the whole genome sequences obtained from genbank database. Each Nucleotide sequence from genbank is used as a query sequence for pair wise alignment.<sup>5</sup>

Pair wise alignment starts between query sequence (lovastatin gene cluster) and the hit sequence (whole genome nucleotide sequences of different fungi), identifies regions of local similarity between the 2 sequences and thus calculates the statistical significance of matches. This offers a convenient approach to calculate the percentage identity.<sup>6</sup>

A pair wise score is calculated for every pair of sequences that are aligned. Percent identity is calculated by multiplying the number of matches in the pair by 100 and dividing by the length of the aligned region, including gaps. Score is calculated by dynamic programming (slow but accurate) or by the method of Wilbur and Lipman (extremely fast but approximate).<sup>7,8</sup> Percent match was calculated using the formula:

Percent Identity = (Matches x 100) / Length of aligned region (with gaps)

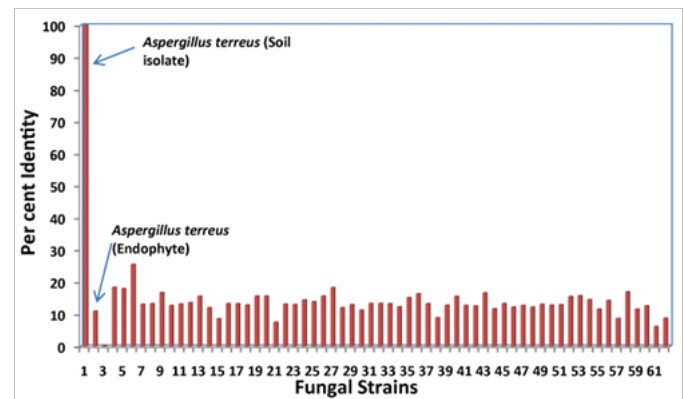
## Results and discussion

The analysis of the whole genome sequence of fungi has proven useful and provided many new insights into secondary metabolism.<sup>9,10</sup> We have previously demonstrated that none of the 54 endophytic fungi isolated from medicinal plants produced any detectable levels of lovastatin even after 10 days of incubation.<sup>4</sup> It was of interest that an endophytic strain of *Aspergillus terreus* did not produce lovastatin while a soil isolate of the same fungus produced significant levels of Lovastatin under identical growth conditions. Therefore, an attempt was made to analyze and compare the genetic makeup of these two isolates of *Aspergillus terreus* and other endophytic fungi in relevance to lovastatin production using bioinformatics tool.

The data set consisted of 62 filamentous fungal whole genome sequences. Sixty one were endophytes including a strain of *Aspergillus terreus* (NT165934.1), and a soil isolate of *Aspergillus terreus* (AH007774.1) which were matched with the available lovastatin gene cluster (AF141924.1 & AF141925.1). All endophytic fungi utilized in this study showed much less homology to lovastatin DNA sequences including an endophytic fungi strain of *Aspergillus terreus* (Figure 1). On the other hand, the soil isolate *Aspergillus terreus* (AH007774.1) showed perfect homology to lovastatin DNA sequences with 100% identity with lovastatin gene cluster sequence (55328/55328) (Table 1, Figure 1). These results are in perfect agreement with our previously published results<sup>4</sup> on the lack of lovastatin production by endophytic fungi and suggest that, the lack of lovastatin production may be due to the absence of lovastatin encoding DNA sequences.

The presence of lovastatin gene cluster in lovastatin producing *Aspergillus terreus* (soil in origin) is well documented.<sup>11,12</sup> However, an endophytic strain of the same organism does not possess

DNA sequences homologous to lovastatin genes. Considering the endophytic relationship with host plant systems, it may be an advantage for a fungal strain not to produce secondary metabolites that may interfere with plant growth promoting characteristics.<sup>13</sup> Thus, there probably was no need for retaining/acquiring lovastatin biosynthetic pathway genes by these endophytic fungi isolated from medicinal plants. This is the first report on comparative analysis of genomic DNA sequence of *Aspergillus terreus* that is both endophytic and non-endophytic in nature. The lovastatin gene cluster includes 18 genes that are involved in lovastatin biosynthetic pathway and, of which LovE and LovF encode for transcriptional regulatory factor and diketide synthase (DKS), respectively.<sup>3,14</sup> Further studies with the endophytic fungi at the molecular level will categorically confirm our previous biochemical<sup>4</sup> and the present bioinformatic study.



**Figure 1** Percent identity for pairwise alignment of different fungi with lovastatin gene cluster DNA sequences.

**Table 1** List of fungi exhibiting their percent identity with lovastatin gene cluster DNA sequences

Sl. no	Fungal organism	NCBI accession number	Percent identity
1	<i>Aspergillus terreus</i> (soil fungus)	AH007774.1	100
<b>Endophytic fungi</b>			
2	<i>Aspergillus terreus</i> NIH2624	NT_165934.1	11.02
3	<i>Colletotrichum gloeosporioides</i> Nara gc5	KB020205.1	17.74
4	<i>Colletotrichum fioriniae</i> PJ7	NW_006889833.1	18.44
5	<i>Colletotrichum higginsianum</i> IMI 349063	CACQ02000033.1	17.98
6	<i>Cucumisattivus</i>	KB722789.1	25.56
7	<i>Glomerella graminicola</i> M1.001	GG697332.1	13.14
8	<i>Glomerella graminicola</i> M1.001	GG697331.1	13.31
9	<i>Glomerella graminicola</i> M1.001	CM001021.1	16.81
10	<i>Ascocoryne sarcoides</i> NRRL 50072	KB205939.1	12.71
11	<i>Ascocoryne sarcoides</i> NRRL 50072	KB205940.1	13.25
12	<i>Ascocoryne sarcoides</i> NRRL 50072	KB205941.1	13.64
13	<i>Leptosphaeria maculans</i>	NW_003533878.1	15.64
14	<i>Pestalotiopsis fici</i> W106-1	KI912109.1	12.09
15	<i>Pestalotiopsis fici</i> W106-1	KI912110.1	8.63
16	<i>Diaporthe longicolla</i> MSPL 10-6	KI547258.1	13.3
17	<i>Diaporthe longicolla</i> MSPL 10-6	KI547259.1	13.3
18	<i>Blumeria graminis</i> f. sp. hordeiDH14	HF938678.1	12.9
19	<i>Blumeria graminis</i> f. sp. hordeiDH14	HF942516.1	15.7

Table Continued...

Sl. no	Fungal organism	NCBI accession number	Percent identity
20	<i>Puccinia striiformis</i> f. sp. triticiCY32	KI515781.1	15.7
21	<i>Puccinia striiformis</i> f. sp. triticiCY32	KI515782.1	7.55
22	<i>Puccinia striiformis</i> f. sp. triticiCY32	KI515783.1	13.2
23	<i>Magnaporthe oryzae</i>	NW_003803360.1	13.06
24	<i>Rhizoctonia solani</i> AG-3 Rh5 IAP strainAG-3	JATN01000001.1	14.49
26	<i>Rhizophagus irregularis</i> DAOM 181602 strain DAOM 197198	KI274019.1	13.92
27	<i>Rhizophagus irregularis</i> DAOM 181602	KI274025.1	15.67
28	<i>Verticillium dahliae</i> JR2	CM001863.1	18.33
29	<i>Ashbya gossypii</i> ATCC 10895	NC_005785.6	12.09
30	<i>Verticillium dahliae</i> JR2	CM001864.1	13.04
31	<i>Ustilago maydis</i> 521	NW_101240.1	11.3
32	<i>Phanerochaete chrysosporium</i> RP-78	DQ097842.1	13.4
33	<i>Sporisorium reilianum</i> SRZ2	FQ311472.1	13.4
34	<i>Nectria haematococca</i> MPVI	GG698896.1	13.33
35	<i>Sclerotinia sclerotiorum</i>	NW_001820835.1	12.34
36	<i>Podospora anserina</i>	NW_001939194.1	42.98
37	<i>Myceliophthora thermophila</i> ATCC 42464	CP003008.1	16.4
38	<i>Claviceps purpurea</i> 20.1	CAGA01000035.1	13.3
39	<i>Cochliobolus sativus</i> ND90Pr	NW_006912048.1	8.92
40	<i>Claviceps purpurea</i>	CAGA01000001.1	12.82
41	<i>Microbotryum violaceum</i> pIA1	GL541643.1	15.62
42	<i>Alternaria brassicicola</i> strain Abra43	AY700092.1	12.799
43	<i>Gaeumannomyces graminis</i> var. tritici	GL385395.1	12.61
44	<i>Alternaria brassicicola</i> brassinin	KC814178.1	16.67
45	<i>Trichoderma atroviride</i> IMI 206040	ABDG02000003.1	11.7
46	<i>Botryotinia fuckeliana</i> B05.10	JH78064.1	13.36
47	<i>Marssonina brunnea</i>	JH921428.1	12.26
48	<i>Botryotinia fuckeliana</i> BcDW1	KB707673.1	12.79
49	<i>Oidiodendron maius</i>	EU386164.1	12.26
50	<i>Bipolaris zeicola</i>	KI964537.1	13.12
51	<i>Fusarium asiaticum</i> strain SCK04	KF765494.1	12.87
52	<i>Marssonina brunnea</i>	NW_006763052.1	13.02
53	<i>Melampsora larici-populina</i>	GL883090.1	15.54
54	<i>Puccinia graminis</i> f. sp. tritici	DS178262.1	15.83
55	<i>Penicillium digitatum</i>	JH993674.1	14.6
56	<i>Blumeria graminis</i>	HF943549.1	11.61
57	<i>Mycosphaerella graminicola</i>	CM001196.1	14.3
58	<i>Cochliobolus heterostrophus</i>	KB445584.1	8.76
59	<i>Dothistroma septosporum</i> NZE10	KB446544.1	16.99
60	<i>Serpula lacrymans</i>	GL945474.1	11.64
61	<i>Heterobasidion irregulare</i>	KI925456.1	12.66
62	<i>Pyrenophora teres</i>	GL537212.1	6.21

## Conclusion

The results of this present analysis provide physical evidence that the apparent inability of endophytic fungi to produce lovastatin secondary metabolite<sup>13</sup> is due solely to the lack of DNA sequences that are identical or homologous to lovastatin gene present in lovastatin producing organism. The lack of production of lovastatin by endophytic fungi is apparently not due to suboptimal physiological conditions or altered regulation of lovastatin biosynthetic pathway enzymes. Owing to the symbiotic association with plant system, the lack of lovastatin homologues in endophytic fungi is of considerable ecological importance and further confirms our hypothesis that endophytic fungi from medicinal plants appear not to produce lovastatin.

## Acknowledgements

The authors would like to thank the financial support provided by the Department of Science & Technology (DST), Govt. of India, vide grant number -DST/SO/FNo.SERB.SR/SO/PS/046/2011.

## Conflict of interest

The author declares no conflict of interest.

## References

1. Sreenivasan A, Shubahgar S, Arvindan R, et al. Microbial production and biomedical applications of lovastatin. *Indian J Pharm Sci.* 2008;70(6):701–709.
2. Slot JC, Rokas A. Horizontal transfer of a large and highly toxic secondary metabolic gene cluster between fungi. *Curr Biol.* 2011;21(2):134–139.
3. Barrios Gonzalez J, Banos JG, Covarrubias AA, et al. Lovastatin biosynthetic genes of *Aspergillus terreus* are expressed differentially in solid-state and in liquid submerged fermentation. *Appl Microbiol Biotechnol.* 2008;79(2):179–186.
4. Praveen VK, Bhargavi SD, Savitha J. Endophytic fungi: A poor candidate for the production of lovastatin. *British Microbiology Research Journal.* 2014;4(12):1511–1520.
5. Bergman NH. *Comparative Genomics.* USA: Humana Press; 2007. 1:1–396.
6. Baxevanis AD, Ouellette BFF. *Bioinformatics: a practical guide to the analysis of genes and proteins.* 3rd ed. USA: Wiley–Interscience; 2005. p. 1–560.
7. Needleman SB, Wunsch CD. A general method applicable to the search for similarities in the amino acid sequence of two proteins. *J Mol Biol.* 1970;48(3):443–453.
8. Wilbur WJ, Lipman DJ. Rapid similarity searches of nucleic acid and protein data banks. *Proc Natl Acad Sci U S A.* 1983;80(3):726–730.
9. Galagan JE, Henn MR, Ma LJ, et al. Genomics of the fungal kingdom: Insights into eukaryotic biology. *Genome Res.* 2005;15(12):1620–1631.
10. Xiao G, Ying SH, Zheng P, et al. Genomic perspectives on the evolution of fungal entomopathogenicity in *Beauveria bassiana*. *Sci Rep.* 2012;2(483):1–10.
11. Subazini TK, Kumar GR. Characterization of lovastatin biosynthetic cluster proteins in *Aspergillus terreus* strain ATCC 20542. *Bioinformation.* 2011;6(7):250–254.
12. Gianoulis TA, Griffin MA, Spakowicz DJ, et al. Genomic analysis of the hydrocarbon-producing, cellulolytic, endophytic fungus *Ascocoryne sarcoides*. *Plos Genetics.* 2012;8(3):e1002558.
13. Aly AH, Debbab A, Kjer J, et al. Fungal endophytes from higher plants: a prolific source of phytochemicals and other bioactive natural products. *Fungal Diversity.* 2010;41(1):1–16.
14. Huang X, Li HM. Cloning and bioinformatic analysis of lovastatin biosynthesis regulatory gene lovE. *Chin Med J (Engl).* 2009;122(15):1800–1805.