Functional annotation of genes associated with colonization and establishment of endophytes in Emilia sonchifolia (LINN.) DC

Sithara K Urumbil
ANILKUMAR MADHAVAN NAIR (drmakumar@gmail.com)
Union Christian College https://orcid.org/0000-0002-8955-0847

Research Article

Keywords: Colonization, gene functional annotation, endophytes, Illumina Miseq, metagenome

Posted Date: May 4th, 2023

DOI: https://doi.org/10.21203/rs.3.rs-2848308/v1

Abstract

Purpose

The entry of rhizosheric bacteria into a plant requires efficient genomic machinery that favours its internalization into the host tissues. We hypothesize that the screening of microbial colonization genes can give an insight into the selection process where some of the soil microbes transform into endophytes in their respective host plants.

Methods

Total DNA was isolated from *Emilia sonchifolia*, PCR amplified and sequenced using Illumina Miseq method. Then metagenome assembly and gene prediction were done followed by gene functional annotation using Nr, KEGG, eggnog and CAZy databases.

Results

Gene functional annotations revealed the presence of genes coding for proteins involved in chemotactic movements such as *cheW, cheY, che* and *CheR*, flagellar and pilin assembly genes, genes associated with biofilm formation viz. *Wza, Wzb, Wzc* and *WcaM*, quorum sensing genes, genes associated with osmoprotectants, host defense etc.

Conclusions

The establishment of endophytes is a complex process and still many lacunae exist in our understanding of the process of internalization and formation of endophytes. This study confirms the presence of genes associated with colonization, internalization and penetration of endophytes in the medicinally important plant *Emilia sonchifolia* and thus adds to our current knowledge in plant-microbe interaction.

Introduction

A mutual beneficial association exists between plants and certain microbes. The diverse microbial community maintains a mutualistic or symbiotic association with plants were introduced as endophytes to the plant research world by De Bary (1866).Their endophytic mode of habitat, symbiotic co-existence, unexplained functional role, host relationship and lifestyle paved way for far-reaching research (Rosenblueth and Martinez-Romero,2006). Majority of the endophytic bacteria reported were from *Proteobacteria* and were considered as soil bacteria (Janssen et al., 2002, Spain et al., 2009).

Endophytes, even in its highly complex interactive life, clearly indicated its effective involvement in the survival, growth and development of the host plant (Potshangbam *et al*., 2017). Endophytes partakes active role in the host development by synthesising bioactive compounds for suppressing pathogenic activities and enhancing plant growth promotion (Mengistu, 2020). Plant root emphasises a selective entry of microbes from soil to root internal tissues since an unbridled entry may leads to pathogenic

invasions. The microbes pass across a series of interactive screening that always act as crucial check points for endophytic entry and establishment (Liu et al., 2017).Therefore Soil microbes develop diverge mechanisms for the internalization and establishment as their journey towards endophytism. Recognision of root exudates, chemotactic movement, internalization, quoram sensing, oxidative stress management and establishment are some of the key points to be monitored in the genome analysis of endophytes to emphasis the journey of soil microbes to endophytes (Compant et al., 2010, Solan and Lebeis, 2015)

The large scale screening with different 'omics' tools made it possible to study the different aspects of plant colonization by soil bacterium. It revealed the molecular basis of interaction between the host cell and the bacterium, the requirement for the bacterium inside the host tissue, and the factors responsible for symbiotic metabolism as a beneficial adaptation. Earlier reports on rhizobia based studies on colonization explained the mechanisms adopted by the soil bacterium for the host identification and penetration to the host internal environment (Desbrosses and Stougaard, 2011). Another face of the metagenome study is that it permits the screening of multiple interactions of plants with beneficial endophytes and pathogens and the endophyte-pathogen interaction also.

Advanced genome technologies shifted the pace of endophytic research which involves the study of multifaceted plant microbe interactions (Khare et al., 2018). The entry of bacterial endophytes in to host cell require numerous systems which helps the endophytes to sense and combat in changing environmental conditions and to endure the induced resistance of the host plant generated through their invasion. The metagenome analysis explained this factor through the presence of different metabolites for ameliorating the oxidative stress and quorum sensing molecules and transporters. In-depth investigation on the genomic data provides the opportunity to assess the biosynthetic machinery of the microbes which were usually not accessible in the culture conditions.

Emilia sonchifolia (Linn.) DC (Asteraceae), widely distributed in India, is used in traditional medicine for curing cuts, wounds, inflammations in throat, ear, eyes etc.(Joshy, 2018, Shen *et al.*,2012). This plant was reported to have anti-inflammatory, cytotoxic and anti-tumorogenic properties in advanced research screenings (Shylesh and Padikkala, 2000, Essien et al., 2009). Biodiversity analysis of endophytic bacterial microflora from *E. sonchifolia* indicated the diversity of bacterial microflora associated with in this medicinally important plant (Urumbil and Anilkumar, 2020). Recent research on the isolated endophytes revealed their role in anti-inflamamtory property of *E. sonchifolia* and was confirmed by in vitro and in vivo studies (Urumbil and Anilkumar, 2021). The in-depth screening of theses microbes reveals effective genomic machineries harboured by these endophytes for the internalization in to the host plant. Since these microbes follows an efficient co-existing relationship with the host plant and involves in the biosynthetic pathways of host plants for the production of novel compounds. Therefore screening of microbial colonization genes in this microbial community will give an insight to the selection of soil microbes and their progressive transformation to endophytes.

Materials and methods

Plant sample collection and surface sterilization

For DNA extraction the whole plants of *E. sonchifolia* at its flowering season (July-October) were collected and pooled from different districts of Kerala, India The plant material was washed with tap water, double distilled, 0.1% mercuric chloride and dipped in 70% ethanol for 60 seconds.

DNA Isolation and Amplification by PCR

DNA was extracted using Purelink genomic DNA extraction and purification kit (invitrogen, Life technologies, USA). The DNA fragments were then separated based on their molecular weight. The concentration of dsDNA was checked by Qubit®4.0 fluorometer. The V3-V4 hyper variable regions from 16S rRNA gene using primers (Forward 5'CCTACGGRRBGCASCAGKVRVGAAT-3', Reverse 5'GGACTACNVGGGTWTCTAA TCC-3') were then amplified.

PCR was carried out for 35 cycles in a Mycycler^{TM} (Bio-Rad, USA) with the initial denaturation at 95ºC for 3min, cyclic denaturation at 98ºC for 20sec, annealing at 54ºC for 30Sec, extension at 72ºC for 1m and ended with a final extension step at 72ºC for 2min. The PCR amplicons were tagged with adapters for creating indexed libraries. Validation of DNA libraries were done by Agilent 2100 Bioanalyzer (Agilent Technologies, Palo Alto, CA, USA) and quantified by Qubit 4.0 Flurometer. DNA libraries were multiplexed and loaded on an Illumina MiSeq instrument according to manufactures instructions (Illumina, 2013, Illumina, San Diego, CA, USA). Sequencing was performed using a 2x250 paired-end (PE) configuration. Image analysis and base calling were performed by the Mi Seq control software (MCS) with in the MiSeq instrument. The original image data were analysed using Bcl2Fastq (V2.17.1.14) and the result stored in FASTQ format. The analysis software used was Cutadapt [(V (1.9.1) V search (1.9.6), Qiime (1.9.1)]. The filtered sequence data were compared with reference database (RDP Gold data base) using UCHIME algorithm to search for chimeric sequences and then chimeric sequences were removed.

Processing and assembly of data

Illumina's built in software decides whether the read to be retained or discarded based on the quality of the first 25 reads. Sequencing data quality optimization was an important step for error reduction before going to downstream analysis. Next generation data quality software Cutadapt (v.1.9.1) was used to trim adaptors, eliminate low quality reads and N-rich reads. The primers and reads with length less than 75bp were also removed in this filtering process. There was chance for the contamination of reads with host sequence, and they were removed with the help of BWA (v0.7.12) software which filtered out host sequences based on host genome.

Metagenome assembly and gene prediction

The clean data was used for whole genome de novo assemblies and was processed using MEGAHIT software (v 1.1.3), with different K-mer (39, 59, 79, 119). After assembling, the scaffold with biggest N50 was selected for further gene prediction analysis. The reads were analysed for coding genes to specify gene predictions of metagenomic or unknown microorganisms using Prodigal (v 3.02). Sequence clustering was done with CD-HIT (v4.5.6) which reduced the redundancy of the predicted gene sequences and clustering of unique gene sequences were processed at 95% identity and 90% coverage level by default. Pre-processed reads were now aligned to non redundant set of genes with the help of Soap Aligner (v2.21) which generated gene abundance or reads coverage of the genes. The number aligned and normalised reads were calculated based on gene length which was used to measure the gene abundance.

Gene functional annotation

Gene functional annotations were predicted by aligning the predicted gene with different databases like Nr database (non-redundant protein database), KEGG pathway database (Kyoto encyclopedia of genes and genomes database), eggNOG (evolutionary geneology of genes: Non Supervised Orthologous Groups, Version 4.0) and CAZy database (Carbohydrate Active enZYmes Database, Diamond Version 0.8.15.77 and BLAST Version 2.2.31+) were used for the database search and alignment to predict the gene functional annotations. Gene annotation resulted from each database was used to categorise relative abundance of different functional categories. The sequenced data were submitted in NCBI Biosample databse SRA with accession no SAMN11617377 and SAMN11616726.

Results

Most of the endophytic microbes were similar to soil microbes and the entry and colonization of these microbes inside the host plant change them to endophytes. So, in the present analysis different genes were identified which contribute to the production of various factors for the entry and establishment of endophytes. The prime factor for the entry of endophyte into the host tissue is the recognition of root exudates from the host plant and chemotactic movement of the microbe towards that direction. In the gene functional annotations, genes coding proteins involved in chemotactic movements like *cheW, cheY, cheB, CheR* were identified (Table 2(1)). A few numbers of genes involved in the assembly of flagella (*flaG, flaB, flaM, flaA* etc.) and pilin (*fimA, fimV, fimU* etc.) were identified and they also play a role in the chemotactic movement and recognition of the host for colonisation (Fig. 1). Some specific components participating in focal adhesion and ECM-receptor interaction (Fig. 2) were recognized and represented as annotated genes in those particular pathways.

Colanic acid like exopolysaccharide was produced by colonic acid biosynthetic gene clusters. Genes like *Wza, Wzb, Wzc, WcaM* were the annotated genes and it indicated the capacity to form biofilm by the endophytes. Colanic acid exporters, protein-tyrosin-phosphatase and colonic acid biosynthesis glycosyl transferase were recognized in the present study (Table 2(2)). Genes responsible for degrading the cell

wall forming polymers like Cellulases, Xylanases and cellobiohydrolases and the plant cell wall softening enzymes, pectate lyases (EC 4.2.2.2.), endoglucanase (EC 3.2.1.4) cellobiohydrolase, cellulase (ec.3.2.1.4) and xylanase (EC 3.2.1.18) (Table 1), were also present. The endophyte microbiome synthesize diverse categories of metabolites which were involved in quorum sensing. Quorum sensing enables the endophyte to sense and respond to the surroundings by the regulation of various genes. Autoinducers, transport systems, proteases and different substrate binding proteins involved in signal transductions during quorum sensing were also identified (Table 2(3)).

Genes participating in the synthetic pathways of osmoprotectants like in the metabolism of taurin, carotenoid synthesis, betaine synthesis and two component systems of signal transductions were identified. Synthesis of trehalose and betaine, with the help of enzymes like trehalose-6-phosphate hydrolase (*ots A,B*), choline dehydrogenase (*bet A*), ATP dependant Potassium(K^+) uptake system genes (*kdp*) which were involved in two component system (TCS) group of signalling pathways were identified in the present analysis. Presence of carotenoid synthesis (an osmoprotectant) was confirmed by the presence of 15-cis-phytoene/All-trans-phytoene synthase (*crtB*).

The endophytic entry into the host cell activates host plant's defence mechanisms. Metagenomic analysis showed the presence of genes for enzymes like *SOD1/SOD2* (superoxide dismutase), *kat* (catalases), *ahpF/C* (peroxiredoxins), and peroxidases (iron dependant and thiol). Along with these enzymes the presence of hydrogen peroxide sensor *OxyR*, multidrug efflux pump (RND family) and osmotically inducible proteins (*osm C/B/E/Y*) (Table 2(4)) showed that the presence and accumulation of reactive oxygen species were managed and reducing the oxidative stress during the establishment of endophyte in the host tissue.

Discussion

The penetration of a soil bacterium into a root tissue of host requires a series of steps to bypass the soil-root interface (Senga et al., 2017). The first step in the process of penetration of bacterium into the host plant is the identification of root exudates of the host and its chemotactic movement towards the root surface (Kandel et al., 2017). The root exudates and rhizodeposits determines or initiates the soil microbe to move towards root surface and these root exudates may be flavanoids, organic acids, different kinds of proteins etc. The soil microbes were chemotactically attracted towards the exudates and they migrate to that direction. Hence it can be inferred that the bacterial genome contain different genes which codes for specific proteins for the recognition and movement. In the present study chemotaxis proteins were recognised from the gene annotations. Studies by Capra and Laub, (2012) showed that *CheA/CheY* are the proteins involved in two-component system of signal transduction and *CheW* is an adaptor protein and also a methyl accepting protein (MCP). In the present study also the *CheA/CheY* (TCS), MCP and adaptor protein *CheA* were recognised. Apart from this some other *Che* proteins like *Z*, and *W* were also noted. The mutation in the *Che* gene will affect the chemotaxis movement and the colonization process, which was proved in the studies with *Herbaspirillum seropedicae* during the colonization on Maize rhizosphere (Balsanelli et al., 2016).

The mobility of bacteria based on chemotaxis can be ascertained in the present study owing to the presence of several genes involved in the biosynthesis, assembly and action of flagella or pili. The role of flagellin protein coding gene *fliC* for the colonization of endophytic *betaproteobacterium Azoarcus* sp. strain BH72 on rice roots were reported (Buschart et al., 2012). It revealed that the colonization of endophyte was arrested when a mutation occurs in *fliC* gene. The interactions with flagellin generate immune responses in host plant, but all the genes responsible for the proper development of flagella was present in the metagenome analysis of rice roots (Sessitsch et al., 2012). These metagenome analyses of endophytic bacteria showed the presence of several genes which imparts its role in the assembly and function of flagella like flagellar hook or associated proteins (*Flg E*, 3*Flg L*, 1*Flg K*), flagellar P-ring proteins (*flgA* and *I*), flagellar L-ring protein (*flgH*) and different flagellar motor switch proteins *FliN/FliY, M,G* (*fliN, Y, M, G*). All the flagellar genes including *fliC* (codes for flagellin). The twitching motility and colonization ability of endophytes were studied based on mutant *pilX* genes and it resulted in reduced colonization (Shidore *et al*., 2012). Here also genes involved in the assembly of type IV pilin (*hof B,C,M,N,O,P,Q*), fimbrial proteins (*fimC, fim I,W,Y*) and pilin (*pilD*) were recognized.

The second major step involved in the colonization was the adhesion of microbes to the root surface. Along with the surface appendages like pili and flagella, other factors were responsible for the adhesion. In the annotations there was the presence of proteins involved in heme utilization or adhesion. *Curli* is an amyloid fibre folding with a major role in adhesion and biofilm formation. They are the protenaceous extra cellular fibres coded by two operons *csg A, B, C* and *csg D, E, F, G* (Barnhart and Chapman, 2006). A homologous operon of *csg* as *afg* was reported from *Salmonella enterica*. Studies based on the inactivated *afg* gene (*afg B*) indicated difficulties in adhesion and attachment to host cells (Barak et al., 2005). In the present study both operons (*csg A,B,C and csg D,E,F,G*) required for the biosynthesis and regulation of *curli* were identified along with another gene *fhaB* which produces another protein hemagglutinin produced by endophytic bacteria for biofilm formation. The role of hemagglutinin like protein in the adhesion and bifilm formation in *Xanthomonas axonopodis* pv. Citri, a phyto-pathogen, was analysed (Gottig et al., 2009). Colonic acid, an exopolysaccharide synthesising genes viz. *Wca E, I, F, J* and *Wza* together with a transcriptional regulator (*mcbR*) was identified in endophytic metagenome analysis of *E. sonchifolia*. It is an important factor for the synthesis and formation of biofilm in *Enterobacteriaceae* (Ratto *et al*., 2006).

Penetration and Internalization

Root surface or rhizoplane attachment followed by its penetration is required for entry of microbes. Some microbes use the natural openings on the plant surface, cracks and wounds to enter in to the plant tissues (Hardoim et al., 2015). In a study by Yi et al., (2017) based on comparative transcriptome analysis, they found that the *Bacillus mycoides* strains shows upregulation of some genes responsible for cell wall modification and degradation when exposed to root exudates. This factor highlights a fact that some endophytes enter inside the plant tissues by degrading the cell wall. The cell wall degrading or modifying factors were responsible for the entry of *Azoarcus* sp. BH72 in to the rice root and the

endoglucanase mutant of the microbe shows decreased penetration and colonization capacity (Reinhold-Hurek et al., 2006). To get information about the genetic potential of endophytic microbes for degrading polysaccharides of plant cell wall, enzyme databases like CAZymes and KEGG were used. From that data it was identified that within the microbiome a number of genes involved in the production of cell wall degrading enzymes like cellulose, xylanase, cellobiohydrolases, pectin acetylesterases etc. were present.

An attempt of entry by any microorganism into the plant tissue activates its immune system and that leads to defence reaction by the generation of reactive oxygen (ROS), reactive nitrogen (RON) and production of phytoalexins (Zeidler et al., 2004). This will create a highly oxidative stress environment for the invading microbes. Hence the endophytic genome will produce different proteins and enzymes for oxidative stress reduction. In this analysis also a wide variety of enzymes and regulators were identified which enable the bacteria to cope up with the oxidative environment. Some of them were superoxide dismutase (Fe-Mn, Cu-Zn families, SOD2/SOD1) catalase peroxidise (*kat G*), peroxidises (*tpx*), peroxiredoxin (*ahpC*) and transcriptional regulator of LysR family (gene *oxyR*) which activates hydrogenperoxide inducible operons. Studies by Alqueres *et al*.,(2013) on rice endophytes makes clear the fact that the reactive oxygen species deactivating enzymes like superoxide dismutase and glutathione reductase are necessary for the colonization of the rice roots by the endophytes. Some endophytic bacteria use different families of multidrug efflux pumps and transporters to survive in this oxidative environment. An efflux pump activity and its role in colonization of *Stenotrophomonas maltophila* D457 in the presence of root exudates were studied. The mutant strain of this particular efflux pump (*smcDEF*) showed reduction in colonization ability (Garcia-Leon *et al*., 2014). Genes coding for different types of multidrug efflux pumps like *vcaM* (ATP-binding cassette, subfamily B, multidrug efflux pump), *mdtP, oprN* (outer membrane protein, multidrug efflux system), *emrA* (membrane fusion protein, multidrug efflux system), *acrB, mexB, adeJ, smeE, mtrD, cmeB* (multidrug efflux pump) were identified in this study.

Osmoprotectants enable bacteria to survive in salt stress conditions. Many such genes were noted in this study which codes for enzymes and proteins involved in the synthesis and transport of osmoprotectants. Choline dehydrogenase (*betA*), betaine aldehyde dehydrogenase (*bet B*), Proline/betaine transporter of MFS family, biosynthetic pathway for arginine and ornithine, glutamine were identified. Proline was already studied and reported as osmoprotectant and their concentration increases in the presence of plant growth promoting rhizobacteria in rice plants (Jha et al., 2011). The accumulation of proline also was reported from plants also to reduce the effect of different stress conditions like temperature, salt etc. (Krasensky and Janak, 2012). The increased level of proline shows an increase in the level of arginine, ornithine and glutamic acid also because they were the precursors in the proline biosynthetic pathway (Ashraf and Foolad, 2007). Different enzymes for the synthesis of Trehalose (osmoprotectant) were identified like trehalose-6-phosphate synthase (*otsA*), trehalose-6-phosphate hydrolase (*treC*) etc. The importance of trehalose in regulating the growth of *Klebsiella* sp. in high NaCl concentration was studied (Woo et al., 2014). A specific pathway for taurine biosynthesis was noted in this study. Along with these osmoprotectants, carotenoid synthesising enzymes were present in the metagenome. Defective phytoene synthase gene (*crtB*, which involve in the carotenoid biosynthetic pathway) in *Pantoea* sp.Yr343 reduced its tolerance to oxidative stress and colonization ability in *Populus deltoides* (Bible et al., 2016).

Conclusion

In the present investigation, the genes associated with penetration, internalization and colonization of endophytes in *Emilia sonchifolia* was identified using metagenome analysis. Functional annotation resulted in the identification of genes associated with chemotactic movement of bacteria towards the root exudates, flagellar assembly, adhesion and biofilm formation, cell wall degrading enzymes, different stress conditions and osmoprotectants. These results indicated that the endophytes were soil microbes and they came in association with the rhizoplane of the host plant which further led to their internalization. The endophytic mode of life cycle enhances the symbiotic relationship between the microbe and plant. The different categories of genes responsible for chemotaxis, adhesion, biofilm formation, cell wall degradation and oxidative stress tolerance identified from the present study shows colonization behaviour and establishment of endophytes in the medicinally important plant *E. sonchifolia*. There are many missing links to our understanding of bacterial colonization that leads to the formation of plant endophytes. In this study, we unravelled the involvement of many such genes linked with the establishment of endophytes in *E. sonchifolia*. Plant -microbe interaction is a complex process that requires extensive research and metagenomic approach can bring many insights into our recent understanding of the role of endophytes in improving plant growth, development, secondary metabolite production, phytoremediation and stress tolerance.

Declarations

Acknowledgements

The first author thanks University Grants Commission, New Delhi, for providing financial assistance by Faculty Development Programme. Both authors gratefully acknowledge the management of Union Christian College for providing infrastructural facilities to carry out the work and also to OmicsGen LifeSciences Pvt Ltd, Kochi, Kerala for metagenome analysis.

Author’s contribution

Both authors equally contributed for the conception, design and execution of the work. The first author performed the experiment, collected data, analyzed and wrote the draft paper and the corresponding author reviewed and finalized the article. Both authors read and approved the final version of the paper.

Data availability

All data associated with this article is available with the corresponding author and can be shared on genuine request.

Competing interest

The authors have no conflict of interest either financially or non-financially.

References

1. Alquéres S, Meneses C, Rouws L et al (2013) The bacterial superoxide dismutase and glutathione reductase are crucial for endophytic colonization of rice roots by *Gluconacetobacter diazotrophicus* PAL5. Mol. Plant Microbe Interact. 26(8):937-945. http://doi.org/10.1094/MPMI-12-12-0286-R
2. Ashraf M, Foolad M (2007) Roles of Glycine, Betaine and Proline in improving plant abiotic stress resistance. Environ Exp Bot 59: 206-216. http://doi.org/10.1016/j.envexpbot. 2005.12.006.
3. Balsanelli E, Tadra-Sfeir MZ, Faoro H, et al (2016) Molecular adaptations of *Herbaspirillum seropedicae* during colonization of the maize rhizosphere. Environ Microbiol 18:2343-2356. https://doi.org/10.1111/1462-2920.12887.
4. Barak JD, Gorski L, Naraghi-Arani P, Charkowski AO (2005) *Salmonella enterica* virulence genes are required for bacterial attachment to plant tissue. App Environ Microbiol 71(10):5685-91. http://doi.org/10.1128/AEM.71.10.5685-5691.2005
5. Barnhart MM, Chapman MR. (2006) Curli biogenesis and function. Annual Rev Microbiol 60:131-47. http://doi.org/10.1146/annurev.micro.60.080805.142106.
6. Bible AN, Fletcher SJ, Pelletier DA (2016) A carotenoid-deficient mutant in *Pantoea* sp. YR343, a bacteria isolated from the rhizosphere of *Populus deltoides*, is defective in root colonization. Front Microbiol 7:491. http://doi.org/:10.3389/fmicb.2016.00491
7. Buschart A, Sachs S, Chen X, Herglotz J, Krause A, Reinhold-Hurek B (2012) Flagella mediate endophytic competence rather than act as MAMPS in rice-*Azoarcus* sp. strain BH72 interactions. Mol Plant Micro Interac 25(2):191-199. http://doi.org/10.1094/ MPMI-05-11-0138.
8. Capra EJ, Laub MT(2012)Evolution of two-component signal transduction systems. Annu Rev Microbiol 66:325- 47. http://doi.org/10.1146/annurev-micro-092611-150039.
9. Compant S, Clement C,Sessitsch, A (2010) Plant growth promoting bacteria in the rhizo- and endosphere of plants: their role, colonization, mechanisms involved and prospects for utilization. Soil Biol Biochem 42:669-678. http://doi.org/10.1016/j.soilbio.2009.11.024
10. De Bary, A (1866) Morphologie der Pilze, Flechten und Myxomyceten Hofmeisterr's Handbook of Physiological Botany. Leipzig: Wilhelm Engelmann.
11. Desbrosses, GJ, Stougaard, J(2011) Root nodulation: a paradigm for how plant-microbe symbiosis influences host developmental pathways. Cell Host Microbe 10, 348–358. http://doi.org/10.1016/j.chom.2011.09.005.
12. Essien GE, Nwidu LL, Nwafor PA (2009) Anti-inflammatory and analgesic potential of methanolic extract of *Emilia Sonchifolia* (Compositae) leaves in rodents. African J Biomedical Res 12, 199-207.
13. García-León G, Hernández A, Hernando-Amado S, Alavi P, Berg G, Martínez, JL (2014) A function of SmeDEF, the major quinolone resistance determinant of *Stenotrophomonas maltophilia*, is the colonization of plant roots. Appl Environ Microbiol 80(15): 4559-4565. http://doi.org/10.1128/AEM.01058-14.

14. Gottig N, Garavaglia BS, Garofalo CG, Orellano EG, Ottado, J (2009) A filamentous hemagglutinin-like protein of *Xanthomonas axonopodis* pv. *Citri*, the phytopathogens responsible for Citrus canker, is involved in bacterial virulence. PLoS ONE. 4:c4358. http://doi.org/10.1371/journal.pone.0004358.
15. Hardoim PR, van Overbeek LS, Berg G, Pirttilä AM, Compant S, Campisano A, Döring M, Sessitsch, A (2015) The hidden world with in plants: ecological evolutionary considerations for defining functioning of microbial endophytes. Microbiol Mol Biol Rev 79:293-320.
16. Janssen PH, Yates PS, Grinton BE, Taylor PM, Sait M (2002) Improved culturability of soil bacteria and isolation in pure culture of novel members of the divisions *Acidobacteria*, *Actinobacteria*, *Proteobacteria*, and *Verrucomicrobia*. App Environ Microbiol 68 (5) 2391-2396. http://doi.org/10.1128/AEM.68.5.2391-2396.2002.
17. Jha Y, Subramanian RB, Patel S(2011) Combination of endophytic and rhizospheric plant growth promoting rhizobacteria in *Oryza sativa* shows higher accumulation of osmoprotectant against saline stress. Acta Physiol Plant 33:797-802. http://doi.org/10.1007/s11738-010-0604-9.
18. Kandel SL, Joubert PM, Doty SL (2017) Bacterial endophyte colonization and distribution within Plants. Microorganisms 5:77. http://doi.org/10.3390/microorganisms 5040077.
19. Khare E, Mishra J, Arora NK (2018)Multifaceted interactions between endophytes and plant: Developments and prospects. Front Microbiol 9:2732. http://doi.org/10.3389/fmicb.2018.02732.
20. Krasensky J, Jonak C (2012) Drought, salt and temperature stress-induced metabolic rearrangements and regulatory networks. J Exp Bot http://doi.org/10.1093/jxb/ err460.
21. Liu H, Carvalhais LC, Crawford M, Singh E, Dennis PG, Pieterse CMJ Schenk PM(2017) Inner plant values: Diversity, colonization and benefits from endophytic bacteria. Front Microbiol http://doi.org/10.3389/fmicb.2017.02552
22. Mengistu A (2020) Endophytes: Colonization, Behaviour and their role in defense mechanism. Int J Microbiol http://doi.org/10.1155/2020/6927219.
23. Potshangbam M, Devi SI, Sahoo D, Strobel GA(2017)Functional cheracterization of endophytic fungal community associated with *Oryza sativa* L. and *Zea mays* L. Front Microbiol 8:325. http://doi.org/10.3389/fmicb.2017.00325.
24. Rättö M, Verhoef R, Suihko ML, Blanco A, Schols AHA, Voragen AGJ, Wilting R, Siika-aho M, Buchert J(2006) Colanic acid is an exopolysaccharide common to many enterobacteria isolated from paper-machine slimes. J Ind Microbiol Biotechnol 33: 359. https://doi.org/10.1007/s10295-005-0064-1.
25. Reinhold-Hurek B, Maes T, Gemmer S, Van Montagu M, Hurek T (2006) An endoglucanase is involved in infection of rice roots by the not-cellulose-metabolizing endophyte *Azoarcus* sp. strain BH72. Mol Plant Microbe Interact 19(2):181-188. http://doi.org/ 10.1094/MPMI-19-0181.
26. Senga RA, Terrazas AS, Balbirnie K, Blank M, Janiak A, Szarejko I, Chmielewska B, Karez J, Morris J, Hedley PE, George TS, Bulgarelli D (2017) Root hair mutations displace the barley rhizosphere microbiota. Front Plant Sci 8:1-15. https://doi.org/10.3389/fpls.2017.01094.
27. Sessitsch A, Hardoim P, Doring J, Weilharter A, Krause A, Woyke T, Mitter B, Hauberg-Lotte L, Friedrich F, Rahalkar M, Hurek T, Sarkar A, Bodrossy L, van Overbeek L, Brar D, van Elsas JD, Reinhold-Hurek

B(2012) Functional characterestics of an endophytic community colonizing rice roots as revealed by metagenomic analysis. MPMI. 25(1):28-36.

28. Shen SM, Shen LG, Lei QF, Si JY, Liu CM, Lu H(2012) Chemical constituents contained in aerial parts of *Emilia sonchifolia.* China J Chinese Materia Medica 37: 3249-3251. http://doi.org/10.11669/cpj.2013.21.008
29. Shidore T, Dinse T, Öhrlein J, Becke A, Reinhold-Hurek B(2012)Transcriptomic analysis of responses to exudates reveals genes required for rhizosphere competence of the endophyte *Azoarcus* sp. strain BH72. Environ Microbiol 14(10):2775-87. Chemical constituents contained in aerial parts of *Emilia sonchifolia.* http://doi.org/10.1111/j.1462-2920.2012.02777.x.
30. Shylesh BS, Padikkala J. (2000) In vitro cytotoxic and antitumor property of Emilia sonchifolia (L.) DC in mice. Journal of Ethnopharmacology, 73, 495-500
31. Solan SS. Lebeis, SL(2015)Exercising influence:distinct biotic interactions shape root microbiomes. Curr Opin Plant Biol 26:32-36.
32. Spain AM, Krumholz LR, Elshahed, MS(2009) Abundance, composition, diversity and novelty of soil *Proteobacteria.* Microbial ecology and functional diversity of natural habitats. The ISME J 992–1000. https://doi.org/10.1038/ismej.2009.43.
33. Urumbil S K Anilkumar M (2020) Diversity analysis of endophytic bacterial microflora in *Emilia sonchifolia* (Linn.) DC on IlluminaMiSeq platforms. J Pure ApplMicrobiol 14(1), 679-687.https://doi.Org/10.22207/jpam.14.1.70
34. Urumbil S K Anilkumar M (2021) Anti-inflamamtory activity of endophytic bacterial isolates from *Emilia sonchifolia* (Linn.) DC. J Ethnopharmacol 281. doi: 10.1016/j.jep.2021.114517.
35. Woo, H.L., Ballo, NR, Hazen et al (2014)Complete genome sequence of the lignin-degrading bacterium *Klebsiella* sp. strain BRL6-2. Stand Genomic Sci 8(9):19. doi: 10.1186/1944-3277-9-19.
36. Yi Y, de Jong A, Frenzel E, Kuipers OP(2017) Comparative transcriptomics of *Bacillus mycoides* Strains in response to Potato-root exudates reveals different genetic adaptation of endophytic and soil isolates. Front Microbiol 8:1487. http://doi.org/ 10.3389/ fmicb. 2017.01487
37. Zeidler D, Zähringer U, Gerber I, Dubery I, Hartung T, Bors W, Hutzler P, Durner J (2004) Innate immunity in *Arabidopsis thaliana*: lipopolysaccharides activate nitric oxide synthase (NOS) and induce defense genes. Proc. Natl. Acad. Sci. USA. 101(44):15811-15816. http://doi.org/10.1073/pnas.0404536101.

Tables

Table- I-CAZy annotations- Cell Wall Degrading Enzymes

Enzyme	CAZy_family	CAZy_activities
Cellulase	GH26	beta-mannanase (EC 3.2.1.78); beta-1,3-xylanase (EC 3.2.1.32)
	GH5	chitosanase (EC 3.2.1.132); Cellulase (EC 3.2.1.4); endo-beta-1,4-xylanase (EC 3.2.1.8); cellulose beta-1,4-cellobiosidase (EC 3.2.1.91) xyloglucan-specific endo-beta-1,4-glucanase (EC 3.2.1.151)
Xylanase	GH11	xylanase (EC 3.2.1.8)
Amylase	GH97	alpha-glucosidase (EC 3.2.1.20); alpha-galactosidase (EC 3.2.1.22)
Cellobio Hydrolases	GH6	endoglucanase (EC 3.2.1.4); cellobiohydrolase (EC 3.2.1.91)
Pectin acetyl esterase	CE12	pectin acetylesterase (EC 3.1.1.-)
maltose-6-phosphate glucosidase	GH4	maltose-6-phosphate glucosidase (EC 3.2.1.122); alpha-glucosidase (EC 3.2.1.20); alpha-galactosidase (EC 3.2.1.22);
thioglucosidase	GH1	thioglucosidase (EC 3.2.1.147);
alpha-glucosidase	GH4	alpha-glucosidase (EC 3.2.1.20);
beta-glucosidase	GH3	beta-glucosidase (EC 3.2.1.21);
alpha-galactosidase	GH4	alpha-galactosidase (EC 3.2.1.22);
beta-galactosidase	GH2	beta-galactosidase (EC 3.2.1.23) ;
alpha-mannosidase	GH47	alpha-mannosidase (EC 3.2.1.113)
beta-mannosidase	GH2	beta-mannosidase (EC 3.2.1.25);
beta-fructo furanosidase	GH68	beta-fructofuranosidase (EC 3.2.1.26);

Table-2-KEGG annotations – Genes associated with plant colonization of endophyte

Table-2(1)-Gene annotations for chemotaxis

Gene_Name	Definition	EC	Pathway
cheBR	two-component system, chemotaxis family, CheB/CheR fusion protein	EC:2.1.1.80 3.1.1.61	ko02020 Two-component system; ko02030 Bacterial chemotaxis;
Mcp	methyl-accepting chemotaxis protein	--	
Chew	purine-binding chemotaxis protein CheW	--	
cheY	two-component system, chemotaxis family, response regulator CheY	--	
cheB	two-component system, chemotaxis family, response regulator CheB	EC:3.1.1.61	
cheR	chemotaxis protein methyltransferase CheR	EC:2.1.1.80	

Table-2(2)-Gene annotations for Biofilm formation/Colanic acid

Gene_Name	Definition	EC	Pathway
wcaE,I,F	putative colanic acid biosynthesis glycosyltransferase WcaE, WcaI	EC:2.4.-.- EC:2.3.1.-	
wcaJ	putative colanic acid biosysnthesis UDP-glucose lipid carrier transferase	--	ko05111 Biofilm formation
mcbR	GntR family transcriptional regulator, colanic acid and biofilm gene transcriptional regulator	--	
Wza	polysaccharide biosynthesis/export protein	--	ko02026 Biofilm formation

Table-2(3)-Gene annotations for Quorum sensing

Gene_Name	Definition	EC	Pathway
luxN	two-component system, autoinducer 1 sensor kinase/phosphatase LuxN	EC:2.7.13.3	ko02020 Two-component system;ko02024 Quorum sensing;
luxS	S-ribosylhomocysteine lyase	EC:4.4.1.21	ko02024 Quorum sensing;
livF	branched-chain amino acid transport system ATP-binding protein	–	ko02010 ABC transporters;ko02024 Quorum sensing;
ftsY	fused signal recognition particle receptor	–	ko02024 Quorum sensing;
vpr	minor extracellular serine protease Vpr	EC:3.4.21.-	
ribA, RIB1	GTP cyclohydrolase II	EC:3.5.4.25	
ABC.SP.S	putative spermidine/putrescine transport system substrate-binding protein	–	

Table-2(4)-Gene annotations for Osmoprotectant/ Oxidative stress reduction

Gene_Name	Definition	EC	Pathway
proP	MFS transporter, MHS family, proline/betaine transporter	--	
proV	glycine betaine/proline transport system ATP-binding protein	EC:3.6.3.32	ko02010 ABC transporters;
betT, betS	choline/glycine/proline betaine transport protein	--	
betB, gbsA	betaine-aldehyde dehydrogenase	EC:1.2.1.8	ko00260 Glycine, serine and threonine metabolism;
TC.BCT	betaine/carnitine transporter, BCCT family	--	
opuD, betL	glycine betaine transporter	--	
betA, CHDH	choline dehydrogenase	EC:1.1.99.1	ko00260 Glycine, serine and threonine metabolism;
otsA	trehalose 6-phosphate synthase	EC:2.4.1.15	ko00500 Starch and sucrose metabolism;
otsB	trehalose 6-phosphate phosphatase	EC:3.1.3.12	ko00500 Starch and sucrose metabolism;
treC	trehalose-6-phosphate hydrolase	EC:3.2.1.93	ko00500 Starch and sucrose metabolism;
treZ, glgZ	maltooligosyltrehalose trehalohydrolase	EC:3.2.1.141	ko00500 Starch and sucrose metabolism;
tauD	taurine dioxygenase	EC:1.14.11.17	ko00430 Taurine and hypotaurine metabolism;
tauC	taurine transport system permease protein	--	ko02010 ABC transporters;
isfD	sulfoacetaldehyde reductase	EC:1.1.1.313	ko00430 Taurine and hypotaurine metabolism;
kdpB	K+-transporting ATPase ATPase B chain	EC:3.6.3.12	ko02020 Two-component system;
KdpE	two-component system, OmpR family, KDP operon response regulator KdpE	--	ko02020 Two-component system;ko02024 Quorum sensing;

KdpA,C,F	K+-transporting ATPase ATPase A,C,F chain	EC:3.6.3.12	ko02020 Two-component system;
kdpD	two-component system, OmpR family, sensor histidine kinase KdpD	EC:2.7.13.3	ko02020 Two-component system;
crtB	15-cis-phytoene/all-trans-phytoene synthase	EC:2.5.1.32 2.5.1.99	ko00906 Carotenoid biosynthesis;
SOD2	superoxide dismutase, Fe-Mn family	EC:1.15.1.1	ko04146 Peroxisome;
SOD1	superoxide dismutase, Cu-Zn family	EC:1.15.1.1	ko04146 Peroxisome;ko04213 Longevity regulating pathway
katG	catalase-peroxidase	EC:1.11.1.21	ko00360 Phenylalanine metabolism
katE, CAT, catB, srpA	catalase	EC:1.11.1.6	ko04146 Peroxisome;ko04211 Longevity regulating pathway;
osmC/B/E/Y	Osmotically/ hyperosmotically inducible proteins	--	
ahpF	alkyl hydroperoxide reductase subunit F	EC:1.6.4.-	
AhpC,	peroxiredoxin (alkyl hydroperoxide reductase subunit C)	EC:1.11.1.15	ko04214 Apoptosis - fly;
Tpx	thiol peroxidase, atypical 2-Cys peroxiredoxin	EC:1.11.1.15	
oxyR	LysR family transcriptional regulator, hydrogen peroxide-inducible genes activator	--	ko02026 Biofilm formation - Escherichia coli;
K06994	putative drug exporter of the RND superfamily	--	

Figures

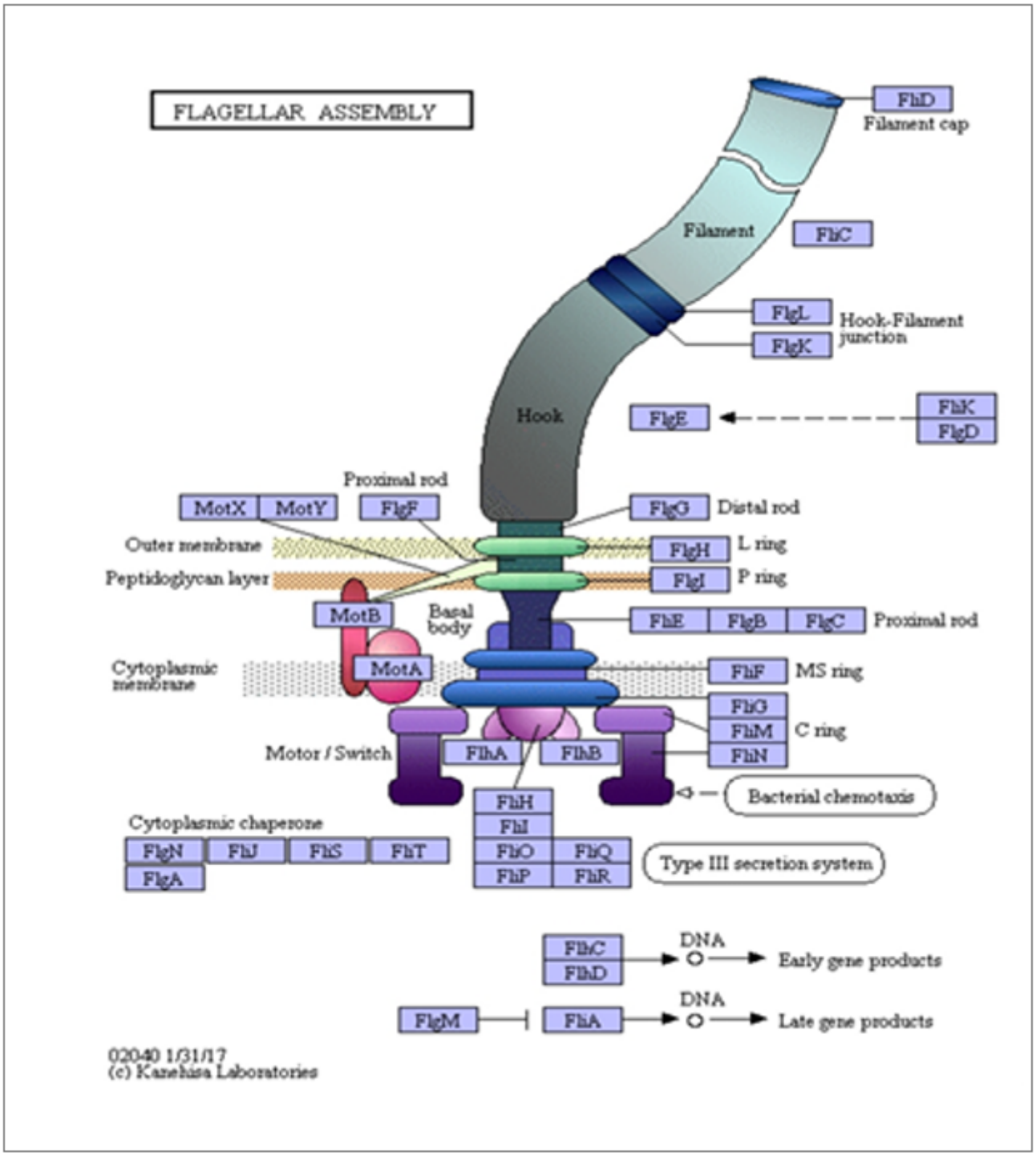

Figure 1

KEGG Annotated Pathway showing genes associated with pilin and flagellar biosynthesis and assembly

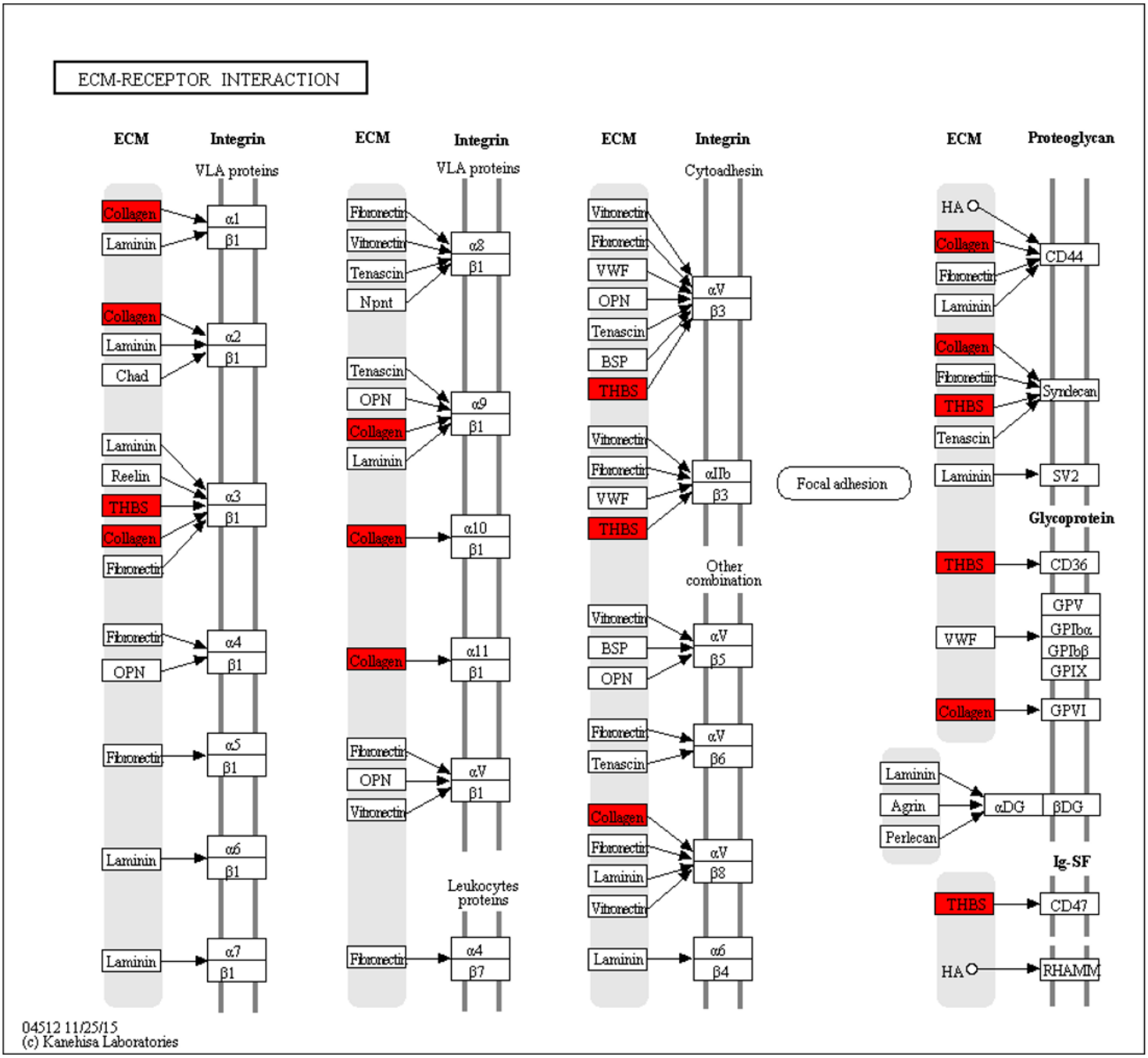

Figure 2

KEGG Annotated Pathway showing genes associated with ECM receptor interactions and foal adhesions