Makita *et al. BMC Genomics* 2018, **19**(Suppl 1):922
DOI 10.1186/s12864-017-4333-y

BMC Genomics

RESEARCH Open Access

Construction of Pará rubber tree genome and multi-transcriptome database accelerates rubber researches

Yuko Makita[1], Mika Kawashima[1], Nyok Sean Lau[1,2], Ahmad Sofiman Othman[2,3] and Minami Matsui[1*]

From 16th International Conference on Bioinformatics (InCoB 2017)
Shenzhen, China. 20-22 September 2017

Abstract

Background: Natural rubber is an economically important material. Currently the Pará rubber tree, *Hevea brasiliensis* is the main commercial source. Little is known about rubber biosynthesis at the molecular level. Next-generation sequencing (NGS) technologies brought draft genomes of three rubber cultivars and a variety of RNA sequencing (RNA-seq) data. However, no current genome or transcriptome databases (DB) are organized by gene.

Results: A gene-oriented database is a valuable support for rubber research. Based on our original draft genome sequence of *H. brasiliensis* RRIM600, we constructed a rubber tree genome and transcriptome DB. Our DB provides genome information including gene functional annotations and multi-transcriptome data of RNA-seq, full-length cDNAs including PacBio Isoform sequencing (Iso-Seq), ESTs and genome wide transcription start sites (TSSs) derived from CAGE technology. Using our original and publically available RNA-seq data, we calculated co-expressed genes for identifying functionally related gene sets and/or genes regulated by the same transcription factor (TF). Users can access multi-transcriptome data through both a gene-oriented web page and a genome browser. For the gene searching system, we provide keyword search, sequence homology search and gene expression search; users can also select their expression threshold easily.

Conclusion: The rubber genome and transcriptome DB provides rubber tree genome sequence and multi-transcriptomics data. This DB is useful for comprehensive understanding of the rubber transcriptome. This will assist both industrial and academic researchers for rubber and economically important close relatives such as *R. communis*, *M. esculenta* and *J. curcas*. The Rubber Transcriptome DB release 2017.03 is accessible at http://matsui-lab.riken.jp/rubber/.

Keywords: *Hevea brasiliensis*, Transcriptome, Gene annotation, Database, Latex biosynthesis, *cis*-1,4-polyisoprene, R-gene

Background

Natural rubber is an indispensable material for many industrial applications such as in tires and medical devices [1]. Although more than 2500 plants produce latex, currently the Pará rubber tree (*Hevea brasiliensis* Muell. Arg.) is the only main commercial source for rubber production [2]. Even compared with petro-chemically synthesized rubber, natural rubber has advantages in adhesion, elasticity and durability.

Natural rubber is produced from specialized differentiated cells called laticifer cells in the outer layer of bark. Rubber latex is composed of rubber serum and rubber particles where *cis*-1,4-polyisoprene biosynthesis occurs. The molecular mechanism of rubber production is not well understood. Additionally, disease-resistance is an important trait to identify for rubber research and breeding. Rubber trees are susceptible to several fungal infections including South American leaf blight (*Microcyclus ulei*) and different cultivars show different sensitivity [3].

* Correspondence: minami@riken.jp
[1]Synthetic Genomics Research Group, Biomass Engineering Research Division, RIKEN Center for Sustainable Resource Science (CSRS), 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
Full list of author information is available at the end of the article

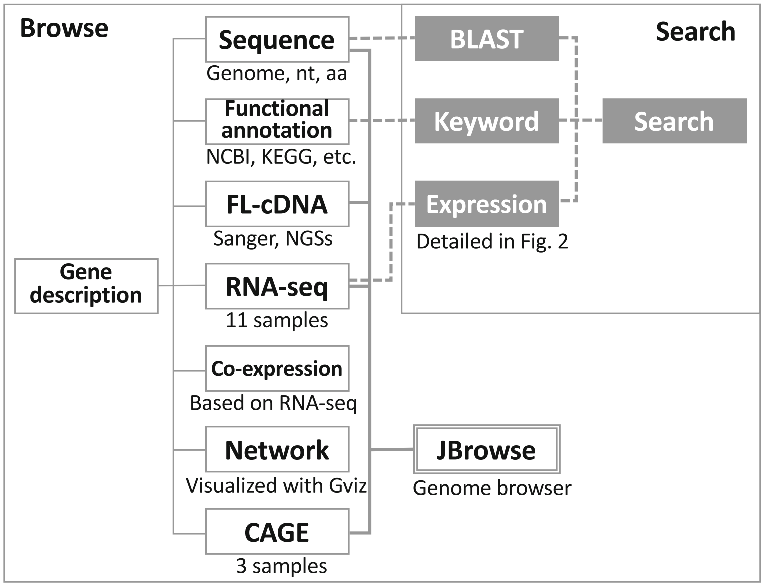


Fig. 1 The structure of rubber transcriptome database. Gene-oriented DB has functional annotation and multi-transcriptome data. Based on 11 RNA-seq data, we calculated co-expressed genes and their networks. All transcriptome data are available from JBrowse, a fast visualization genome browser. Three types of search systems are also available. See Fig. 2 for the details of the gene expression search

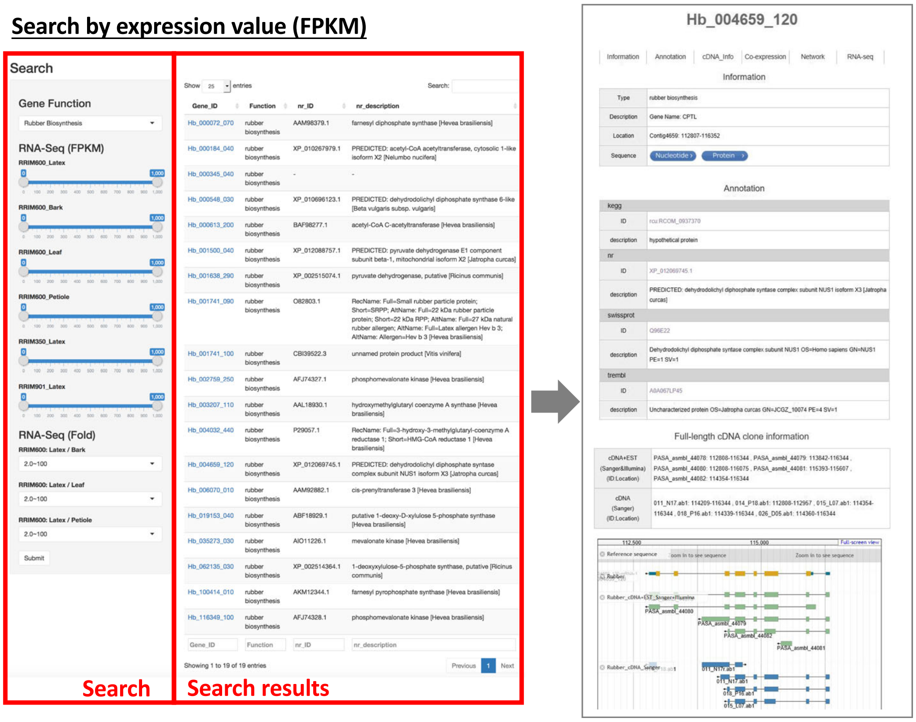
To accelerate molecular biological research in *H. brasiliensis*, we first determined its draft genome sequence and annotated 84,443 protein-coding genes [4]. After the rubber tree genome was determined, transcriptome data is important to identify gene expression level and precise gene structure, such as transcription start sites (TSS) and isoforms. Currently, RNA-seq is the most widely used transcriptome technology, providing gene expression levels in many conditions and tissues. For the rubber tree, latex where *cis*-1,4-polyisoprene biosynthesis occurs is the key tissue to understand the mechanism of rubber production. Many researchers determined gene expression of latex in different cultivars or conditions [5–9]. Although RNA-seq is a powerful tool to know gene expression, it is difficult to predict full-length splice isoforms and TSSs accurately. For better understanding of transcription in rubber tree, we constructed full-length cDNA libraries and determined their sequence with Sanger and Illumina [10]. Pootakham et al. released Pacific Biosciences (PacBio) Isoform sequencing (Iso-Seq), single-molecule real-time long-read isoform sequencing in BPM24 cultivar [11].

When we predict 5′-end with EST and/or RNA-seq, we tend to predict the longest TSSs instead of major expressed TSSs [12]. To know major expressed TSS in different tissues, we applied cap analysis gene

Table 1 Data source and number of annotated genes of each database or experiment

Category	Contents	Number of annotated rubber genes	URL of the data source	Reference
Genome and proteins	Draft genome	84,443	http://matsui-lab.riken.jp/rubber/	[4]
Functional annotation	KEGG	61,453	http://www.genome.jp/	[15]
	NCBI Protein	64,585	https://www.ncbi.nlm.nih.gov/protein/	[16]
	Swiss-Prot	33,553	http://www.uniprot.org/uniprot/	[17]
	TrEMBLE	63,053		
	GO	35,247	http://www.geneontology.org/	[18]
Transcriptome	FL-cDNA	7704	http://matsui-lab.riken.jp/rubber/	[10]
	RNA-seq	42,614		[10] [23]
	CAGE	21,168		[4]
	ESTs	23,790[a]	http://www4a.biotec.or.th/rubber/ http://scarecrow.fmrp.usp.br/heveabr/	[19–21]
	Iso-Seq	17,668	http://www4a.biotec.or.th/rubber/	[11]

[a]Number of annotated rubber genes was calculated using three EST data sources and our original FL-cDNA



Search by expression value (FPKM)

Search

Gene Function: Rubber Biosynthesis

RNA-Seq (FPKM)

RRIM600_Latex: 0 – 1,000

RRIM600_Bark: 0 – 1,000

RRIM600_Leaf: 0 – 1,000

RRIM600_Petiole: 0 – 1,000

RRIM350_Latex: 0 – 1,000

RRIM901_Latex: 0 – 1,000

RNA-Seq (Fold)

RRIM600: Latex / Bark: 2.0~100

RRIM600: Latex / Leaf: 2.0~100

RRIM600: Latex / Petiole: 2.0~100

Submit

Show 25 entries

Search:

Gene_ID	Function	nr_ID	nr_description
Hb_000072_070	rubber biosynthesis	AAM98379.1	farnesyl diphosphate synthase [Hevea brasiliensis]
Hb_000184_040	rubber biosynthesis	XP_010267979.1	PREDICTED: acetyl-CoA acetyltransferase, cytosolic 1-like isoform X2 [Nelumbo nucifera]
Hb_000345_040	rubber biosynthesis	-	-
Hb_000548_030	rubber biosynthesis	XP_010696123.1	PREDICTED: dehydrodolichyl diphosphate synthase 6-like [Beta vulgaris subsp. vulgaris]
Hb_000613_200	rubber biosynthesis	BAF98277.1	acetyl-CoA C-acetyltransferase [Hevea brasiliensis]
Hb_001500_040	rubber biosynthesis	XP_012088757.1	PREDICTED: pyruvate dehydrogenase E1 component subunit beta-1, mitochondrial isoform X2 [Jatropha curcas]
Hb_001638_290	rubber biosynthesis	XP_002515074.1	pyruvate dehydrogenase, putative [Ricinus communis]
Hb_001741_090	rubber biosynthesis	O82803.1	RecName: Full=Small rubber particle protein; Short=SRPP; AltName: Full=22 kDa rubber particle protein; Short=22 kDa RPP; AltName: Full=27 kDa natural rubber allergen; AltName: Full=Latex allergen Hev b 3; AltName: Allergen=Hev b 3 [Hevea brasiliensis]
Hb_001741_100	rubber biosynthesis	CBI39522.3	unnamed protein product [Vitis vinifera]
Hb_002759_250	rubber biosynthesis	AFJ74327.1	phosphomevalonate kinase [Hevea brasiliensis]
Hb_003207_110	rubber biosynthesis	AAL18930.1	hydroxymethylglutaryl coenzyme A synthase [Hevea brasiliensis]
Hb_004032_440	rubber biosynthesis	P29057.1	RecName: Full=3-hydroxy-3-methylglutaryl-coenzyme A reductase 1; Short=HMG-CoA reductase 1 [Hevea brasiliensis]
Hb_004659_120	rubber biosynthesis	XP_012069745.1	PREDICTED: dehydrodolichyl diphosphate syntase complex subunit NUS1 isoform X3 [Jatropha curcas]
Hb_006070_010	rubber biosynthesis	AAM92882.1	cis-prenyltransferase 3 [Hevea brasiliensis]
Hb_019153_040	rubber biosynthesis	ABF18929.1	putative 1-deoxy-D-xylulose 5-phosphate synthase [Hevea brasiliensis]
Hb_035273_030	rubber biosynthesis	AIO11226.1	mevalonate kinase [Hevea brasiliensis]
Hb_062135_030	rubber biosynthesis	XP_002514364.1	1-deoxyxylulose-5-phosphate synthase, putative [Ricinus communis]
Hb_100414_010	rubber biosynthesis	AKM12344.1	farnesyl pyrophosphate synthase [Hevea brasiliensis]
Hb_116349_100	rubber biosynthesis	AFJ74328.1	phosphomevalonate kinase [Hevea brasiliensis]

Showing 1 to 19 of 19 entries

Search | Search results

Hb_004659_120

Information | Annotation | cDNA_Info | Co-expression | Network | RNA-seq

Information

Type	rubber biosynthesis
Description	Gene Name: CPTL
Location	Contig4659: 112807-116352
Sequence	Nucleotide › Protein ›

Annotation

kegg	
ID	rcu:RCOM_0937370
description	hypothetical protein
nr	
ID	XP_012069745.1
description	PREDICTED: dehydrodolichyl diphosphate syntase complex subunit NUS1 isoform X3 [Jatropha curcas]
swissprot	
ID	Q96E22
description	Dehydrodolichyl diphosphate syntase complex subunit NUS1 OS=Homo sapiens GN=NUS1 PE=1 SV=1
trembl	
ID	A0A067LP45
description	Uncharacterized protein OS=Jatropha curcas GN=JCGZ_10074 PE=4 SV=1

Full-length cDNA clone information

cDNA+EST (Sanger&Illumina) (ID:Location)	PASA_asmbl_44078: 112808-116344 , PASA_asmbl_44079: 113842-116344 , PASA_asmbl_44080: 112808-116075 , PASA_asmbl_44081: 115393-115607 , PASA_asmbl_44082: 114354-116344
cDNA (Sanger) (ID:Location)	011_N17.ab1: 114209-116344 , 014_P18.ab1: 112808-112957 , 015_L07.ab1: 114354-116344 , 018_P16.ab1: 114339-116344 , 026_D05.ab1: 114360-116344

Fig. 2 Screen capture and explanation of the expression search function. Users' can easily select the expression threshold of each samples with FPKM value or Fold-change value

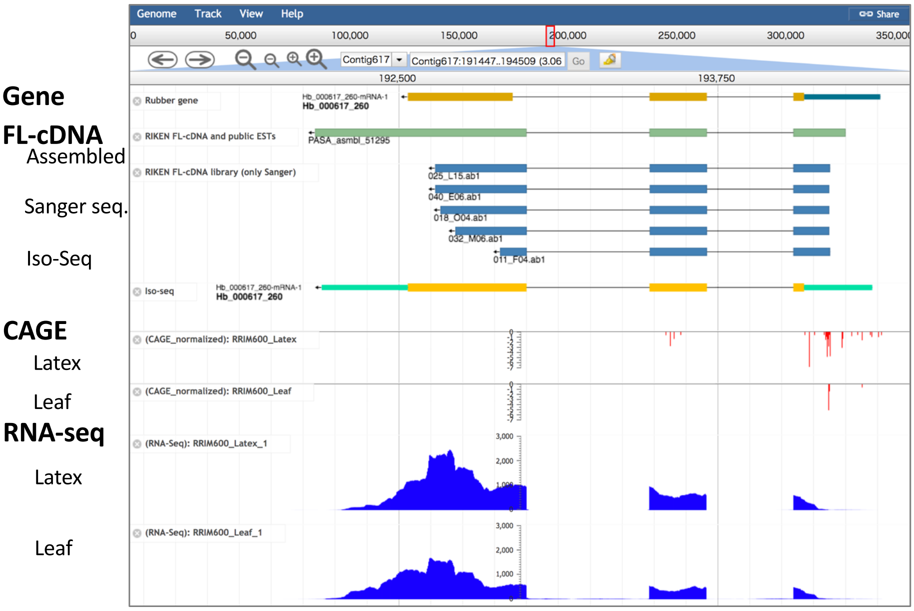
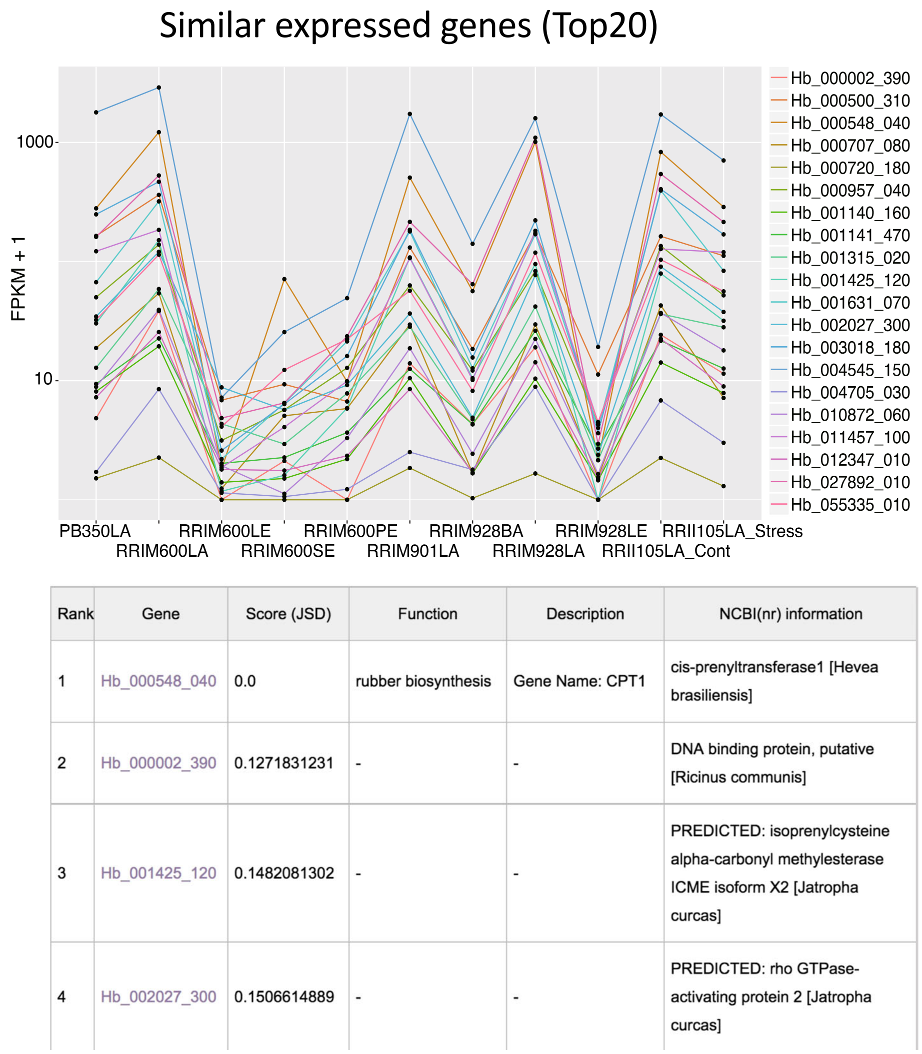


Fig. 3 Gene, FL-cDNA, RNA-seq and CAGE data are visualized on the genome browser. We use the JBrowse fast and embeddable genome browser. Sequence information is also available from JBrowse



Rank	Gene	Score (JSD)	Function	Description	NCBI(nr) information
1	Hb_000548_040	0.0	rubber biosynthesis	Gene Name: CPT1	cis-prenyltransferase1 [Hevea brasiliensis]
2	Hb_000002_390	0.1271831231	-	-	DNA binding protein, putative [Ricinus communis]
3	Hb_001425_120	0.1482081302	-	-	PREDICTED: isoprenylcysteine alpha-carbonyl methylesterase ICME isoform X2 [Jatropha curcas]
4	Hb_002027_300	0.1506614889	-	-	PREDICTED: rho GTPase-activating protein 2 [Jatropha curcas]

Fig. 4 Top 20 of similar expressed genes of CPT1 (*cis*-prenyltransferase 1). The expression patterns (above) were calculated and visualized using R library cummeRbund. The bottom table shows functional annotation and scores (Jensen–Shannon divergence of similar expressed 10 genes

expression (CAGE) method that captures the 5′ end of the transcribed and capped mRNAs. CAGE provided us with a genome wide single base-pair resolution map of TSSs [13].

One of the main difficulties for gene annotation in non-model plants is the number of genes with unknown functions. In our case, 22,991 genes are rubber specific and functionally unknown and 33,213 genes have homologous sequences but are still functionally unknown. To overcome the problem, we carried out co-expressed analysis to predict functionally related gene groups. We can expect genes regulated by the same transcription factor (TF) or genes involved in the same biological pathway to show a similar expression pattern. There are many co-expression databases in plant [12, 14]. In this DB, we show the top 20 similarly expressed genes in rubber.

Construction, content and utility

General structure and content of rubber Transcriptome DB

Based on our original draft genome sequence of *H. barasiliensis* RRIM 600, we predicted 84,443 protein-coding genes [4]. Users can access functional annotation and three kinds of transcriptome data through web pages organized by gene. The basic structure of our DB is summarized in Fig. 1. Users can access functional annotation and links to original databases; NCBI protein, KEGG, UniProt (both Swiss-Prot and TrEMBLE) and Gene Ontology [15–18]. We also provide multi-transcriptome data of our original full length cDNA (FL-cDNA), RNA-seq and CAGE, and publicly available ESTs, RNA-seq and Iso-Seq (Table 1) [4, 10, 11, 19–22]. Since FL-cDNA sequences help to improve predicted gene structures, we constructed two FL-cDNA libraries and obtained c.a. 20,000 clones [10]. Our DB also

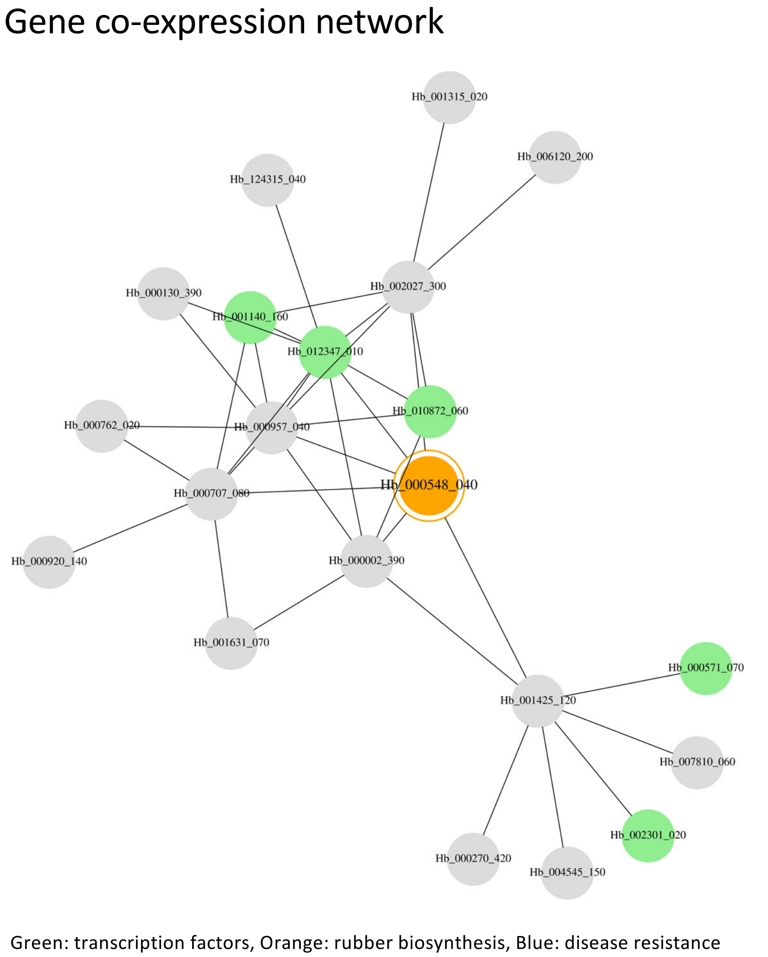


Fig. 5 Two steps of co-expressed gene network of CPT1. Transcription factors are highlighted with green

includes public Iso-Seq and ESTs to capture various isoforms [11, 19–21]. Using RNA-seq technology, we aimed to reveal expression features of natural rubber biosynthetic genes. We previously obtained latex and non-latex (leaf, bark and petiole) in RRIM 600 genome cultivar and latex of other cultivars (RRIM 900 and PB350). Additionally, we downloaded latex and non-latex (bark and leaf) in RRIM 928 and latex under normal and stressed condition in RRII 105 cultivars. All data were re-analyzed with the same protocol [10] and descriptions of all transcriptomic data were summarized in additional file 1 (Table S1). All our original data can be downloaded from our web site.

Data search system

We provide three types of searching systems: keyword search, blast homology search and gene expression search. Since functional annotation of rubber genes is still limited, keyword search and homology search are not enough. We prepared gene expression search so that users can easily change gene expression threshold for samples and retrieve gene sets such as latex specific, for example. In our system, users can select gene expression value of FPKM (Fragments Per Kilobase Million) or select fold-change and obtain gene sets according users' specific demands (Fig. 2).

Genome browser

We visualized our multi-transcriptome data in a genome browser (Fig. 3). We made links for the genome browser on each gene web page. Users can visualize eleven RNA-seq, two FL-cDNA and three CAGE data and compare tissue specificity easily.

Co-expressed genes and their network view

Genes that are regulated by the same transcription factors or that work on the same biological pathways often show similar expression patterns. To suggest candidate genes that may have similar biological function with a user's genes of interest, we calculated the top 20 similar expressed genes using our original and publically available 11 samples (RRIM600 cultivar: latex, bark, leaf and petiole, RRIM901 cultivar: latex, PB350 cultivar: latex [10], RRIM928 cultivar: latex, bark and leaf (SRP022257), RRIM105 latex: control latex and stressed latex (SRP017288)). To visualize the co-expressed gene network (Fig. 4), we used Gviz and highlighted transcription factors (Fig. 5).

Rubber biosynthetic pathway search

The natural rubber biosynthetic pathway is not fully understood, especially regarding the polymerization process with isopentenyl diphosphates (IPPs). To clear known factors and their paralogous genes, we prepared a natural rubber biosynthetic pathway (Fig. 6). Users can easily understand the whole picture of the pathway and access gene information.

Implementation of this web site

Rubber genome & transcriptome databased are currently running on Linux (Ubuntu 16.04.3 LTS) with the following environments: Perl (ver. 5.22.1), PHP (ver. 7.0.18), Python 2.7 and Apache HTTP server (ver. 2.4.18). As a relational database management system, we set up the MySQL (ver. 5.7.18). The co-expression network graph is drawn with Gviz (ver. 2.30.1-19.el7). Fast data access on the genome browser was implemented by JBrowse (ver. 1.12.3) [23]. Gene expression search was built with the Shiny package in R (https://shiny.rstudio.com/). We configured Shiny server to run multiple Shiny processes. Gene description pages are generated as static web pages.

Discussion

Recently, RNA-seq is a powerful tool and the most widely spread transcriptome technology. However, RNA-seq is not suitable for predicting TSSs and full-length splice isoforms accurately. To obtain an entire picture of the rubber transcriptome, it is important to integrate multi-transcriptome data. In this DB, we integrated three types of transcriptome data: RNA-seq, FL-cDNA and CAGE. FL-cDNA provide us precise gene structures including findings of novel genes and novel alternative isoforms. To enrich the quality and quantity of splicing information it is necessary to know the correct protein sequence and estimate the protein function. Iso-Seq, PacBio single molecule long-read, is a powerful technology for this purpose. To know the genome-wide TSSs, we carried out CAGE and obtained precise TSSs in latex, leaf and bark. With CAGE data, we can observe tissue-dependent alternative TSSs, multiple TSSs in a gene and variations of TSS, such as strictly determined TSS in a base or others. In latex biosynthetic pathway, 1-deoxy-D-xylulose-5-phosphate synthase gene and pyruvate dehydrogenase gene showed the different TSS patterns between samples. Precise TSSs assist to find consensus sequences or motif sequences for transcription factors and their regulations.

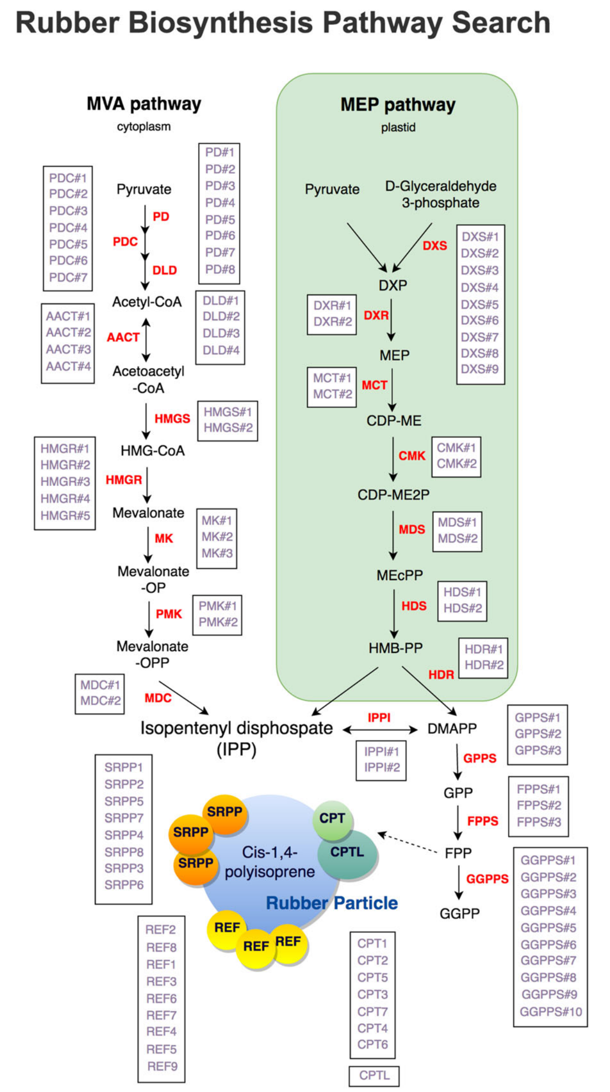


Fig. 6 Rubber biosynthesis pathway search. Each gene is clickable and user can easily access to detailed gene page

As a next target, we plan to expand *cis*-elements that are predicted after co-expressed gene set.

Conclusions

We have developed the rubber tree genome and transcriptome database, a comprehensive and searchable database of economically important plant, *H. brasiliensis*. To assist researchers and breeders in using the *H. brasiliensis* genome, we prepared a simple and user-friendly interface. We believe this database assists both industrial and academic researchers for rubber and important industrial close relatives such as *M. esculenta, R. communis and J. curcas*.

Additional files

Additional file 1: Table S1. Detailed information on the transcriptome data. (PDF 12 kb)

Abbreviations

CAGE: Cap analysis of gene expression; DB: Database; FL-cDNA: Full-length cDNA; Iso-Seq: Isoform sequencing; NGS: Next-generation sequencing; RNA-seq: RNA sequencing; TF: Transcription factor; TSS: Transcription start site

Acknowledgements

We thank David Gifford (RIKEN) for his constructive comments on our manuscript. We thank Yoshiki Mochizuki (RIKEN) and Akihiro Matsushima (RIKEN) for their technical advices on our web site.

Funding

This research is conducted under the research program of RIKEN Biomass Engineering. Publication charges for this article have been funded to YM by Incentive Research Project of RIKEN.

Availability of data and materials

All our original data can be downloaded from our web site. The raw data was registered DDBJ/EMBL/Genbank BioProject under accession PRJDB4387.

About this supplement

This article has been published as part of BMC Genomics Volume 19 Supplement 1, 2018: 16th International Conference on Bioinformatics (InCoB 2017): Genomics. The full contents of the supplement are available online at https://bmcgenomics.biomedcentral.com/articles/supplements/volume-19-supplement-1.

Authors' contributions

MK constructed the database. NSL carried out gene functional annotations. MK and YM analyzed transcriptome data. YM and MM conceived the study. YM prepared the manuscript. MM revised the manuscript. All authors read and approved the final manuscript.

Ethics approval and consent to participate

Not applicable.

Consent for publication

Not applicable.

Competing interests

The authors declare that they have no competing interests.

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Author details

[1]Synthetic Genomics Research Group, Biomass Engineering Research Division, RIKEN Center for Sustainable Resource Science (CSRS), 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan. [2]Centre for Chemical Biology, Universiti Sains Malaysia, 11900 Bayan Lepas, Penang, Malaysia. [3]School of Biological Sciences, Universiti Sains Malaysia, 11800 Minden, Penang, Malaysia.

Published: 19 January 2018

Reference

1. Mooibroek H, Cornish K. Alternative sources of natural rubber. Appl Microbiol Biotechnol. 2000;53(4):355–65.
2. van Beilen JB, Poirier Y. Establishment of new crops for the production of natural rubber. Trends Biotechnol. 2007;25(11):522–9.
3. Lieberei R. South American leaf blight of the rubber tree (Hevea spp.): new steps in plant domestication using physiological features and molecular markers. Ann Bot. 2007;100(6):1125–42.
4. Lau NS, Makita Y, Kawashima M, Taylor TD, Kondo S, Othman AS, Shu-Chien AC, Matsui M. The rubber tree genome shows expansion of gene family associated with rubber biosynthesis. Sci Rep. 2016;6:28594.
5. Tang C, Yang M, Fang Y, Luo Y, Gao S, Xiao X, An Z, Zhou B, Zhang B, Tan X, et al. The rubber tree genome reveals new insights into rubber production and species adaptation. Nat Plants. 2016;2(6):16073.
6. Chao J, Chen Y, Wu S, Tian WM. Comparative transcriptome analysis of latex from rubber tree clone CATAS8-79 and PR107. Genom Data. 2015;5:120–1.
7. Chow KS, Ghazali AK, Hoh CC, Mohd-Zainuddin Z. RNA sequencing read depth requirement for optimal transcriptome coverage in *Hevea brasiliensis*. BMC Res Notes. 2014;7:69.
8. Liu JP, Zhuang YF, Guo XL, Li YJ. Molecular mechanism of ethylene stimulation of latex yield in rubber tree (*Hevea brasiliensis*) revealed by de novo sequencing and transcriptome analysis. BMC Genomics. 2016;17:257.
9. Li D, Wang X, Deng Z, Liu H, Yang H, He G. Transcriptome analyses reveal molecular mechanism underlying tapping panel dryness of rubber tree (*Hevea brasiliensis*). Sci Rep. 2016;6:23540.
10. Makita Y, Ng KK, Veera Singham G, Kawashima M, Hirakawa H, Sato S, Othman AS, Matsui M. Large-scale collection of full-length cDNA and transcriptome analysis in *Hevea brasiliensis*. DNA Res. 2017;24(2):159–67.
11. Pootakham W, Sonthirod C, Naktang C, Ruang-Areerate P, Yoocha T, Sangsrakru D, Theerawattanasuk K, Rattanawong R, Lekawipat N, Tangphatsornruang S. De novo hybrid assembly of the rubber tree genome reveals evidence of paleotetraploidy in Hevea species. Sci Rep. 2017;7:41457.
12. Makita Y, Shimada S, Kawashima M, Kondou-Kuriyama T, Toyoda T, Matsui M. MOROKOSHI: transcriptome database in Sorghum bicolor. Plant Cell Physiol. 2015;56(1):e6.
13. Shiraki T, Kondo S, Katayama S, Waki K, Kasukawa T, Kawaji H, Kodzius R, Watahiki A, Nakamura M, Arakawa T, et al. Cap analysis gene expression for high-throughput analysis of transcriptional starting point and identification of promoter usage. Proc Natl Acad Sci U S A. 2003;100(26):15776–81.
14. Aoki Y, Okamura Y, Tadaka S, Kinoshita K, Obayashi T. ATTED-II in 2016: a plant Coexpression database towards lineage-specific Coexpression. Plant Cell Physiol. 2016;57(1):e5.
15. Kanehisa M, Furumichi M, Tanabe M, Sato Y, Morishima K. KEGG: new perspectives on genomes, pathways, diseases and drugs. Nucleic Acids Res. 2017;45(D1):D353–61.
16. Clark K, Karsch-Mizrachi I, Lipman DJ, Ostell J, Sayers EW. GenBank. Nucleic Acids Res. 2016;44(D1):D67–72.
17. Consortium GO. Gene ontology consortium: going forward. Nucleic Acids Res. 2015;43(Database issue):D1049–56.
18. The UniProt Consortium. UniProt: the universal protein knowledgebase. Nucleic Acids Res. 2017;45(D1):D158–69.
19. Chow KS, Wan KL, Isa MN, Bahari A, Tan SH, Harikrishna K, Yeang HY. Insights into rubber biosynthesis from transcriptome analysis of *Hevea brasiliensis* latex. J Exp Bot. 2007;58(10):2429–40.
20. Salgado LR, Koop DM, Pinheiro DG, Rivallan R, Le Guen V, Nicolás MF, de Almeida LG, Rocha VR, Magalhães M, Gerber AL, et al. De novo transcriptome analysis of *Hevea brasiliensis* tissues by RNA-seq and screening for molecular markers. BMC Genomics. 2014;15:236.
21. Triwitayakorn K, Chatkulkawin P, Kanjanawattanawong S, Sraphet S, Yoocha T, Sangsrakru D, Chanprasert J, Ngamphiw C, Jomchai N, Therawattanasuk K, et al. Transcriptome sequencing of *Hevea brasiliensis* for development of microsatellite markers and construction of a genetic linkage map. DNA Res. 2011;18(6):471–82.
22. Kodama Y, Shumway M, Leinonen R, Collaboration INSD. The sequence read archive: explosive growth of sequencing data. Nucleic Acids Res. 2012;40(Database issue):D54–6.
23. Buels R, Yao E, Diesh CM, Hayes RD, Munoz-Torres M, Helt G, Goodstein DM, Elsik CG, Lewis SE, Stein L, et al. JBrowse: a dynamic web platform for genome visualization and analysis. Genome Biol. 2016;17:66.