Universidade Federal de São Paulo (UNIFESP)

I Curso de Verão em Biotecnologia

PROCURANDO PROMOTORES ESPECÍFICOS DE COLMO PARA APLICAÇÕES BIOTECNOLÓGICAS

Eng. Amanda Fanelli
Eng. Maycow Berbert
Prof.Dr. Michael dos Santos Brito

São José dos Campos, 15 de fevereiro de 2019

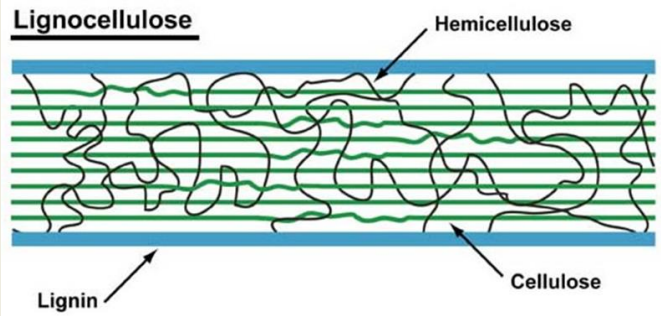
O uso da biomassa de gramíneas (cana, milho, arroz, sorgo)

Material lignocelulósico

Palha

Bagaço

Parede celular



Etanol 2G

Bioprodutos

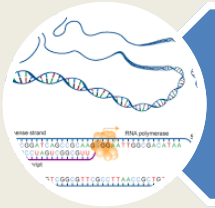
Alimentação animal

Problema: Estrutura complexa da parede!
Como acessar os polímeros?

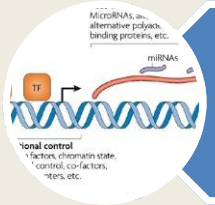
Manipulação do DNA e engenharia metabólica

Obter uma planta com lignina/hemicelulose com estruturas diferentes ou quantidades diferentes

Conhecer a função dos genes (Parede celular)



Controlar a expressão gênica



Promotores e fatores de transcrição

Análise de genes diferencialmente expressos em sorgo

- Sorghum bicolor

(modelo para cana-de-açúcar)

- Plataforma GEO

Usar o R para criar uma ferramenta de mineração de dados

NCBI

GEO Gene Expression Omnibus

HOME | SEARCH | SITE MAP | GEO Publications | FAQ | MIAME | Email GEO

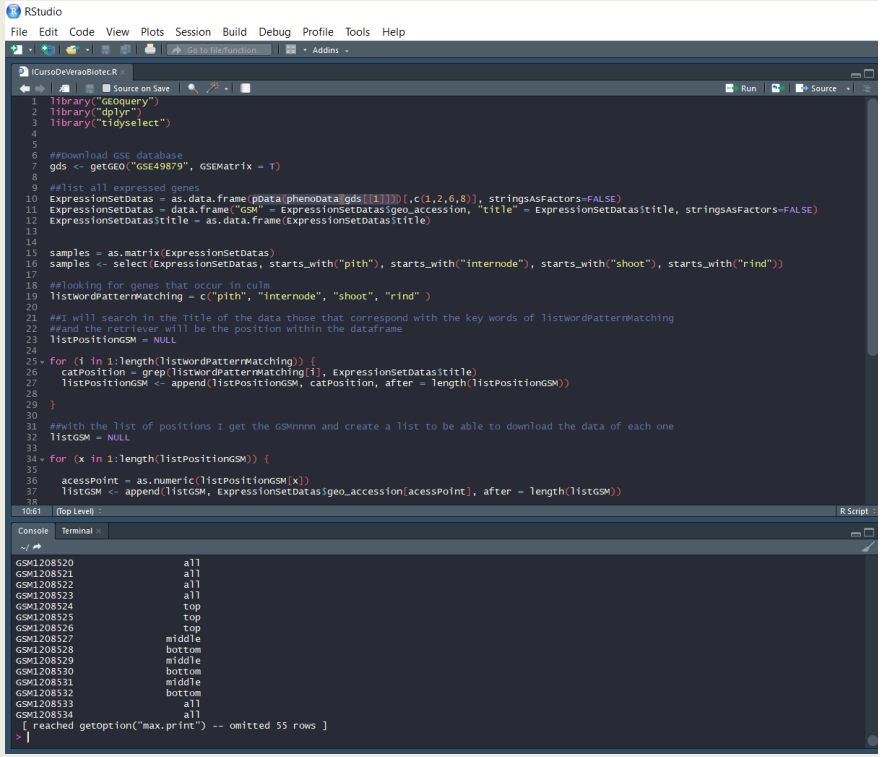
NCBI > GEO > **Accession Display** Not logged in | Login

Scope: Self Format: HTML Amount: Quick GEO accession: GSE49879 GO

Series GSE49879 Query DataSets for GSE49879

Status	Public on Dec 31, 2013
Title	Expression data from vegetative tissues of grain, sweet and bioenergy sorghums
Platform organisms	Sorghum bicolor; Saccharum hybrid cultivar; Zea mays subsp. mays
Sample organism	Sorghum bicolor
Experiment type	Expression profiling by array Non-coding RNA profiling by array
Summary	We developed a commercially available whole-transcriptome sorghum microarray (Sorgh-WTa520972F) and generated this dataset to identify tissue and genotype-specific expression patterns for all identified Sorghum bicolor exons and UTRs. The genechip contains 1,026,373 probes covering 149,182 exons (27,577 genes) across the Sorghum bicolor nuclear, chloroplast and mitochondrial genomes. Specific probesets were also included for putative non-coding RNAs that may play a role in gene regulation (e.g., microRNAs), and confirmed functional small RNAs in related species (corn and sugarcane) were also included in our array design.
Overall design	78 samples were analyzed from four different tissue types (shoot, seedling, leaves and stem), two dissected stem tissues (pith and rind) and six diverse genotypes (PI455230, Atlas, PI152611, AR2400, R159, and Fremont)
Contributor(s)	Shakoor N, Feltus A, Kresovich S
Citation(s)	Shakoor N, Nair R, Crasta O, Morris G et al. A Sorghum bicolor expression atlas reveals dynamic genotype-specific expression profiles for vegetative tissues of grain, sweet and bioenergy sorghums. *BMC Plant Biol* 2014 Jan 23;14:35. PMID: 24456189
Submission date	Aug 14, 2013
Last update date	Jan 28, 2014
Contact name	Nadia Shakoor
Organization name	Chromatin, Inc.
Street address	10 South LaSalle Street, Suite 2100

Quando visualizei a estrutura de dados do GEO pela primeira vez no R



https://github.com/MBerbert/ICursoVeraoBiotecUnifespSJC/blob/master/README.md

Análises dos genes usando a plataforma do site

- Genótipos de biomassa PI455230, PI152611,AR2400
- Dois grupos
  - Colmo: Pith, rind, Shoot, internode
  - Outros: Leaf,root, seedling root

Genes diferencialmente expressos com função de parede celular ou fator de transcrição

Proteins	Domains
Sb02g036010.1	PF03552 AT1G02730.1 ATCSLD5 ATCSLD5; 1,4-beta-D-xylan synthase/ cellulose synthase
Sb02g002200.1	PF03790,PF03791,PF03789 PTHR11850 KOG0773 AT4G08150.1 KNAT1 KNAT1 (KNOTTED-LIKE FROM ARABIDOPSIS THALIANA); transcription factor
Sb02g036000.1	PF03552 AT1G02730.1 ATCSLD5 ATCSLD5; 1,4-beta-D-xylan synthase/ cellulose synthase
Sb01g009460.1	PF03790,PF03791,PF03789 PTHR11850 KOG0773 AT4G08150.1 KNAT1 KNAT1 (KNOTTED-LIKE FROM ARABIDOPSIS THALIANA); transcription factor
Sb04g033930.1	PF00319,PF01486 PTHR11945 KOG0014 AT2G22540.1 SVP SVP (SHORT VEGETATIVE PHASE); transcription factor/ translation repressor, nucleic acid binding
Sb01g021720.1	AT1G08050.1 zinc finger (C3HC4-type RING finger) family protein
Sb01g006790.1	PF03790,PF03791,PF03789 PTHR11850 KOG0773 AT4G08150.1 KNAT1 KNAT1 (KNOTTED-LIKE FROM ARABIDOPSIS THALIANA); transcription factor
Sb04g008670.1	PF00249 AT2G38300.1 DNA binding / transcription factor
Sb02g012640.1	AT5G60710.1 zinc finger (C3HC4-type RING finger) family protein
Sb06g000920.1	PF00092 PTHR10166 AT5G60710.1 zinc finger (C3HC4-type RING finger) family protein
Sb10g021360.1	PF00249 AT2G38300.1 DNA binding / transcription factor
Sb07g005070.1	PF01370 PTHR10366 KOG1502 AT2G33600.1 cinnamoyl-CoA reductase family
Sb09g002080.1	PF00847 AT2G28550.1 RAP2.7 RAP2.7 (RELATED TO AP2.7); DNA binding / transcription factor
Sb03g003640.1	PF03106 AT4G39410.1 WRKY13 WRKY13; transcription factor
Sb05g002940.1	PF03479 AT4G12080.1 DNA-binding family protein
Sb03g003370.1	PF03106 AT5G15130.1 WRKY72 WRKY72; transcription factor
Sb03g009840.1	PF01370 PTHR10366 KOG1502 AT1G15950.1 CCR1 CCR1 (CINNAMOYL COA REDUCTASE 1); cinnamoyl-CoA reductase
Sb01g046040.1	PF05678 AT3G56710.1 SIB1 SIB1 (SIGMA FACTOR BINDING PROTEIN 1); binding / protein binding
Sb09g026100.1	PF02309 AT3G16500.1 PAP1 PAP1 (PHYTOCHROME-ASSOCIATED PROTEIN 1); transcription factor
Sb04g027540.1	PF00249 PTHR10641 KOG0048 K09422 AT1G34670.1 AtMYB93 AtMYB93 (myb domain protein 93); DNA binding / transcription factor
Sb06g001430.1	PF08240,PF00107 PTHR11695 KOG0023 K00095 AT4G39330.1 CAD9 CAD9 (CINNAMYL ALCOHOL DEHYDROGENASE 9); binding / catalytic/ oxidoreductase/ zinc ion binding
Sb01g009480.1	PF03790,PF03791,PF03789 PTHR11850 KOG0773 AT4G08150.1 KNAT1 KNAT1 (KNOTTED-LIKE FROM ARABIDOPSIS THALIANA); transcription factor
Sb08g001850.1	PF00931 AT3G07040.1 RPM1 RPM1 (RESISTANCE TO P. SYRINGAE PV MACULICOLA 1); nucleotide binding / protein binding
Sb01g014310.1	PF02365 AT3G18400.1 anac058 anac058 (Arabidopsis NAC domain containing protein 58); transcription factor
Sb07g024550.1	PF12171,PF00096 PTHR11389 KOG3576 AT2G02080.1 AtIDD4 AtIDD4 (Arabidopsis thaliana Indeterminate(ID)-Domain 4); transcription factor

Análise dos promotores

Genes DE

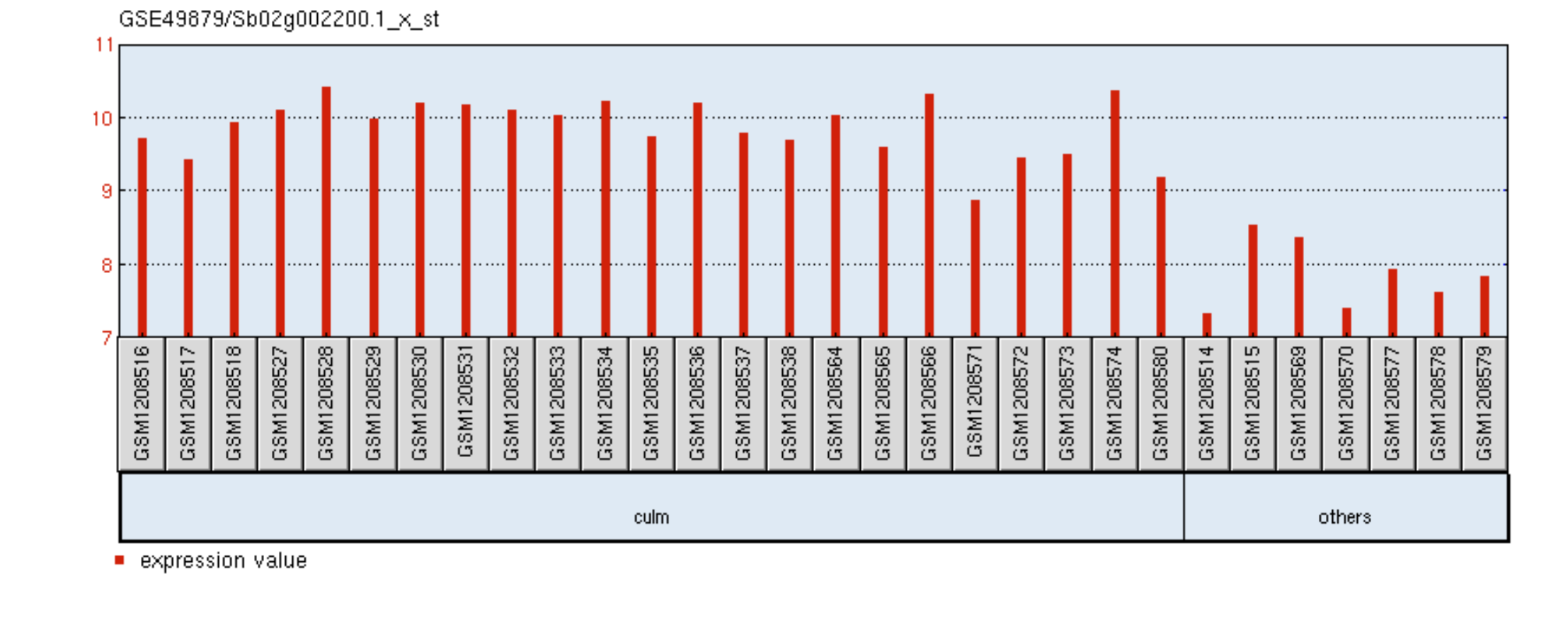
- Promotores(genes)
- Rede de regulação(fatores de transcrição)

PlantCARE Cis-Acting Regulatory Element

STRING

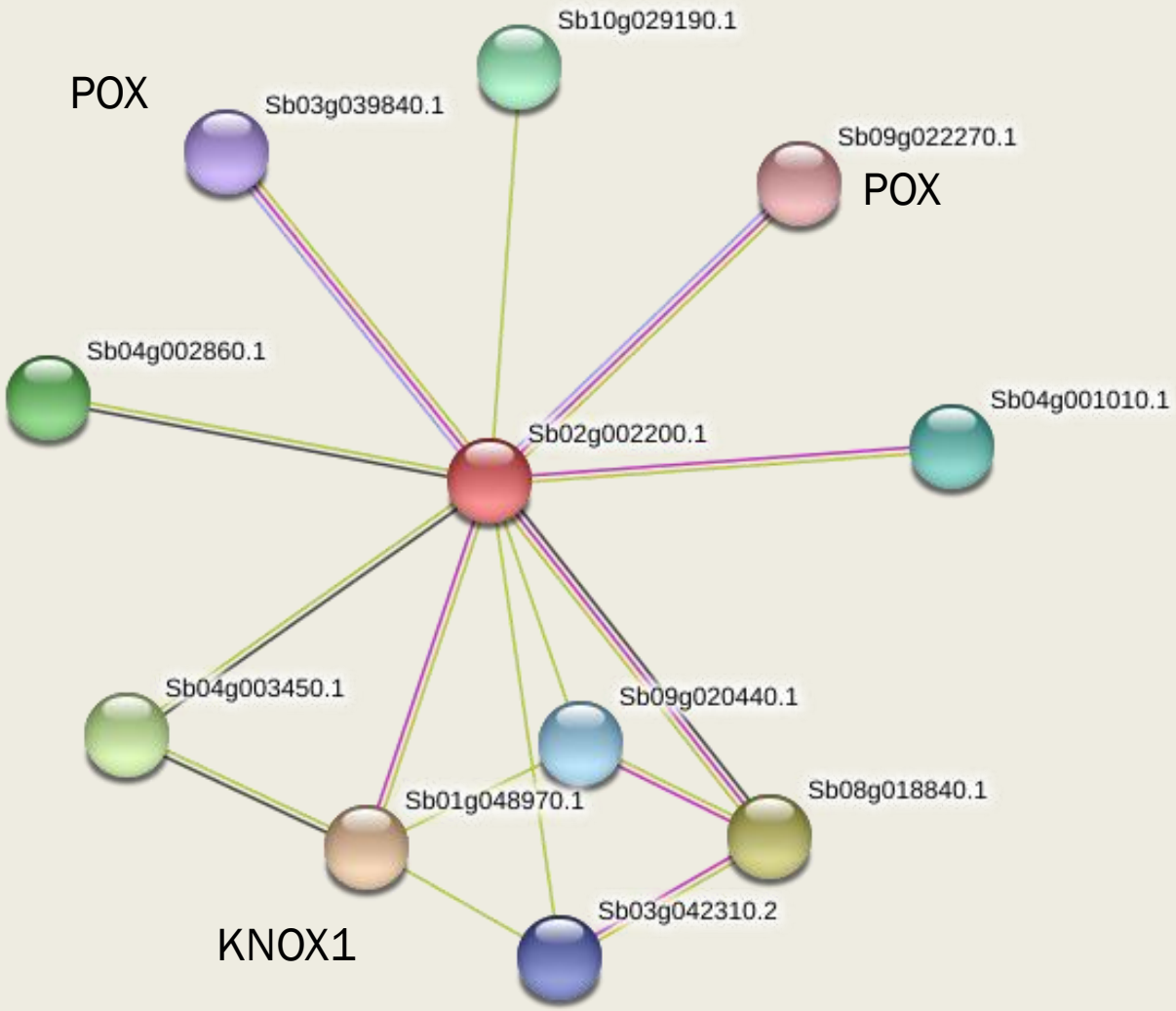
KNAT1 Sb02g002201.1

Fator de transcrição KNAT1:
classI Knox gene relacionado com o desenvolvimento de fibras (Em Arabidopsis)

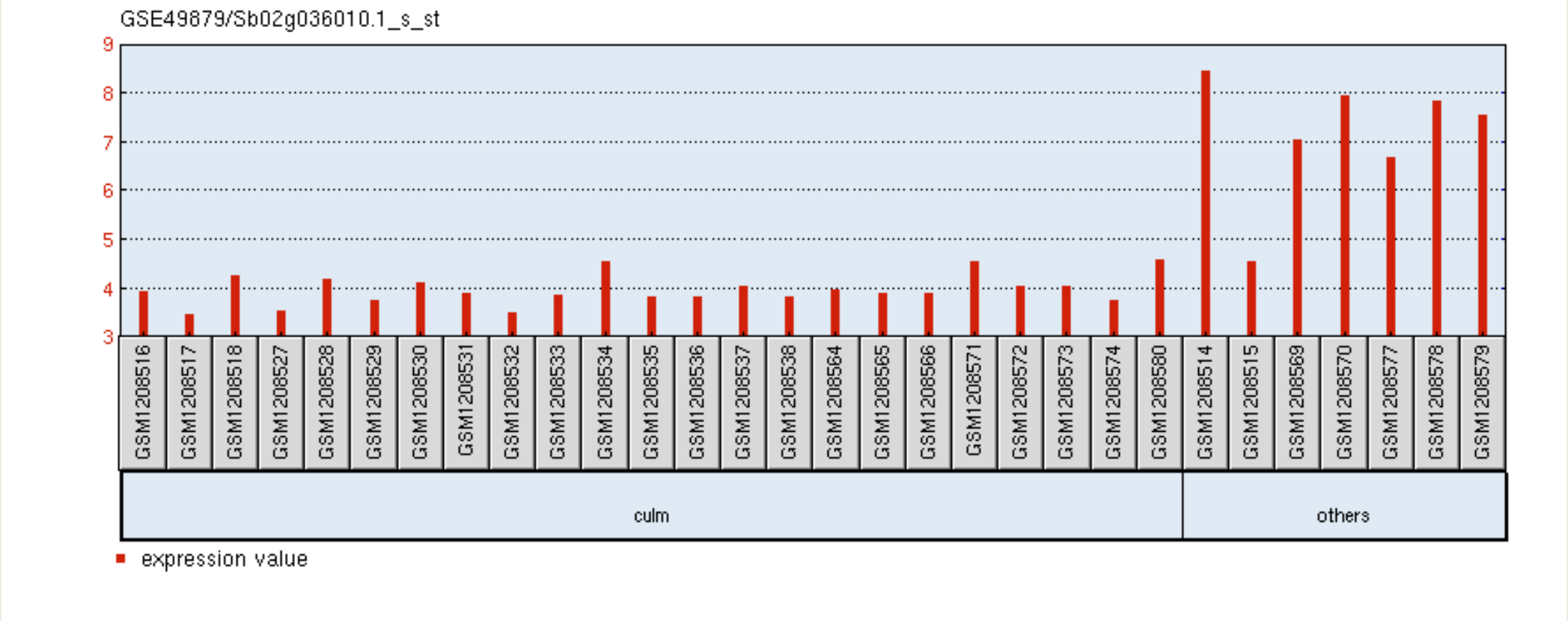


KNAT1 Sb02g002200.1

Análise da rede de regulação



Ces Sb02g036010.1

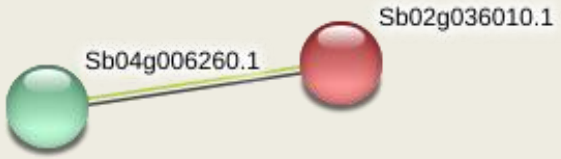


Elementos cis

Ces Sb02g036010.1

- ABRE
- ARE
- AT~TATA-box
- Box 4
- CAAT-box
- CAT-box
- CCAAT-box
- CCGTCC motif
- CCGTCC-box
- CGTCA-motif
- DRE1
- G-Box
- G-box
- GA-motif
- GARE-motif
- GATA-motif
- HD-Zip 1
- MBS
- MYB
- MYB recognition site
- MYC
- Myb
- Myb-binding site
- STRE
- TATA-box
- TC-rich repeats
- TGACG-motif
- Unnamed__1
- Unnamed__4

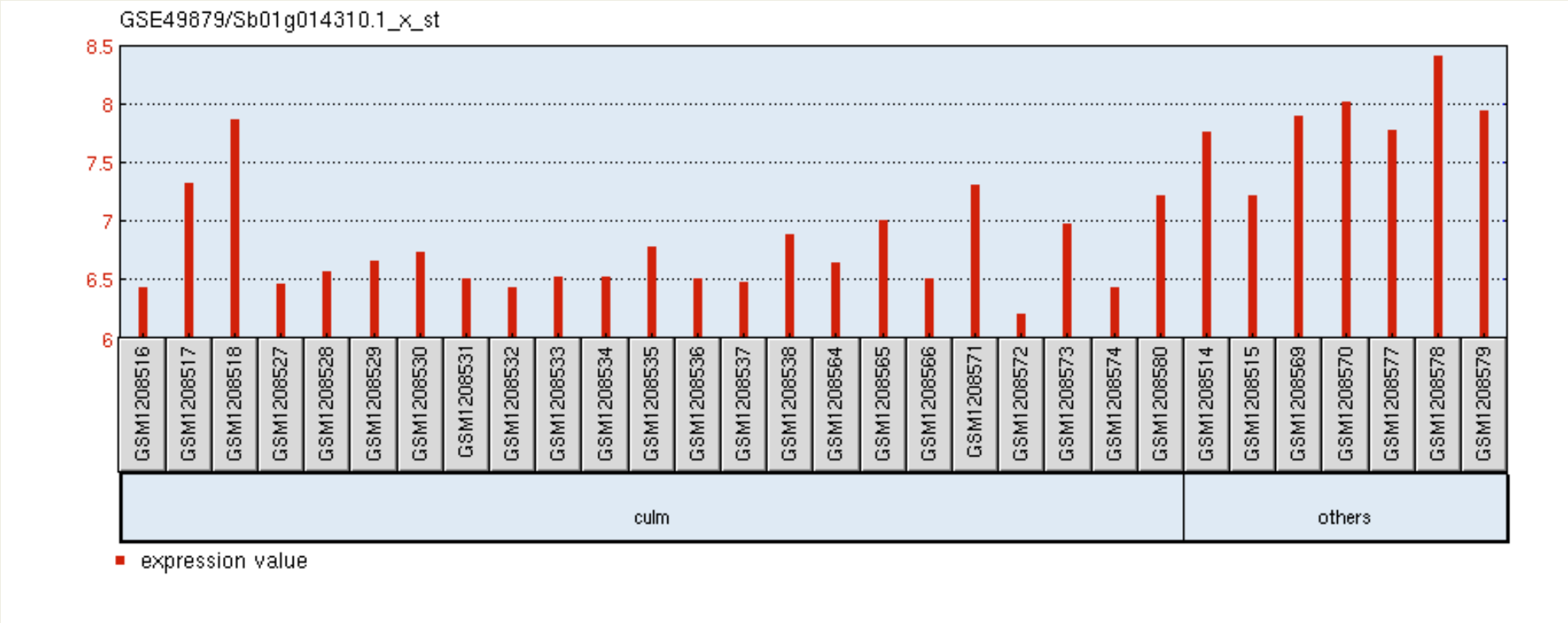
Ces Sb02g036010.1
Análise da rede de regulação



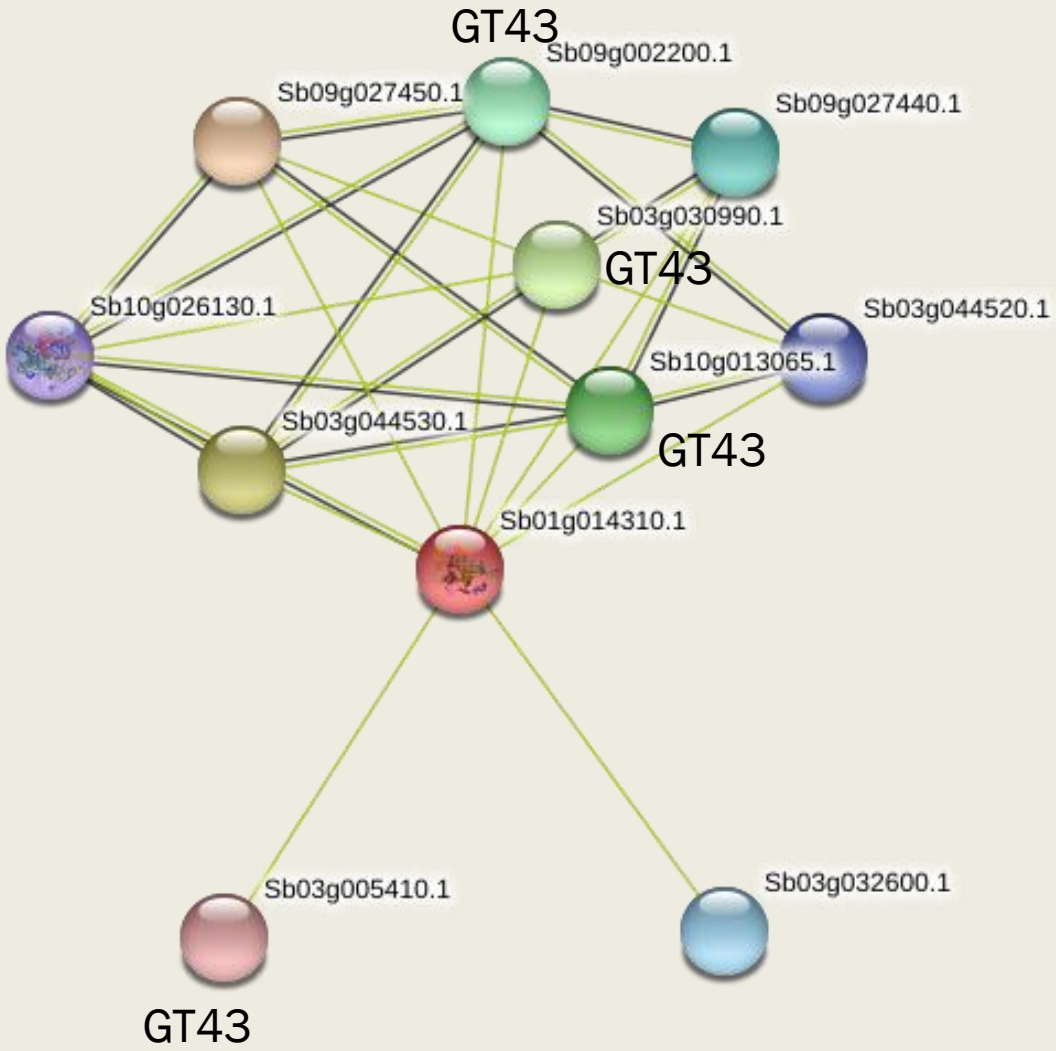
GT2
(CSL)

Também envolvida na síntese de celulose e hemicelulose

NAC Sb01g014310.1



NAC Sb01g014310.1
Análise da rede de regulação



Genes de síntese
de hemicelulose

Conclusão

- Métodos de mineração de dados
- Estudar o promotor dos genes KNAT1
- Genes de celulose e lignina, fatores de transcrição possivelmente envolvidos na hemicelulose (NAC)

Agradecimentos

- ICT-Unifesp
- Prof.Dr. Michael Brito
- Alunos do laboratório (Alice e Marinara)