

VIỆN KHOA HỌC NÔNG NGHIỆP VIỆT NAM
VIỆN DI TRUYỀN NÔNG NGHIỆP

HỘI NGHỊ KHOA HỌC NĂM 2017

“Ứng dụng tin sinh học trong nghiên cứu xác định các họ gen liên quan đến tính chống chịu ở cây trồng”

Báo cáo viên : NCS. Chu Đức Hà

NỘI DUNG BÁO CÁO

ĐẶT VẤN ĐỀ

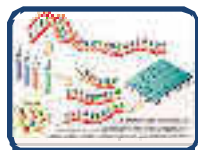
MỘT SỐ KẾT QUẢ ĐÃ ĐẠT ĐƯỢC

ĐỊNH HƯỚNG NGHIÊN CỨU

KẾT LUẬN VÀ ĐỀ NGHỊ

ĐẶT VẤN ĐỀ

Tin sinh học



GENOMICS

Giải trình tự DNA, lập bản đồ di truyền, công nghệ DNA tái tổ hợp...



PROTEOMICS

Xác định protein, định lượng protein, biến đổi sau dịch mã...



METABOLOMICS

Nghiên cứu về các hợp chất chuyển hóa, hormones và các phân tử tín hiệu trong tế bào...



PHENOMICS

Đánh giá đặc điểm hình thái, thiết lập mối liên kết giữa di truyền, epigenetic và yếu tố môi trường...



INOMICS

Xác định yếu tố và tương tác giữa chúng, quá trình điều hòa hóa sinh học...

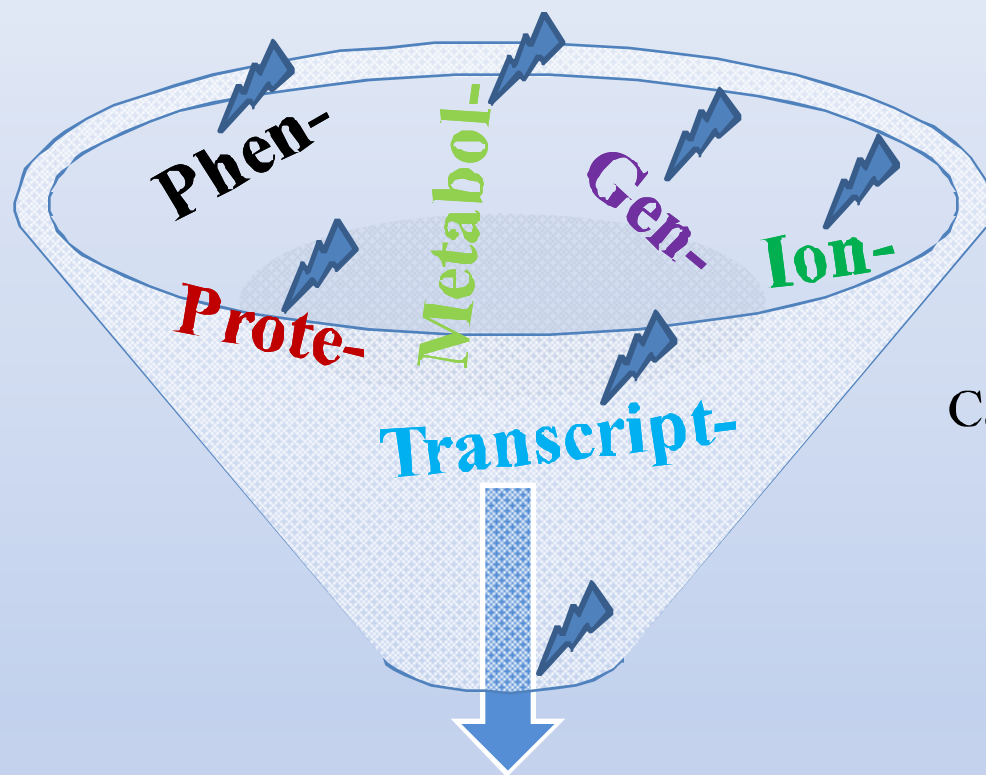


TRANSCRIPTOMICS

Giải trình tự RNA, đánh giá biểu hiện, điều hòa phiên mã...

Tăng cường tính chống chịu điều kiện ngoại cảnh bất lợi ở cây trồng.

Hình 1. Các lĩnh vực -omics quan trọng được sử dụng trong nghiên cứu trên cây trồng (theo Henry T Nguyen và cs., 2014)



Các phương pháp tiếp cận “-OMICS”



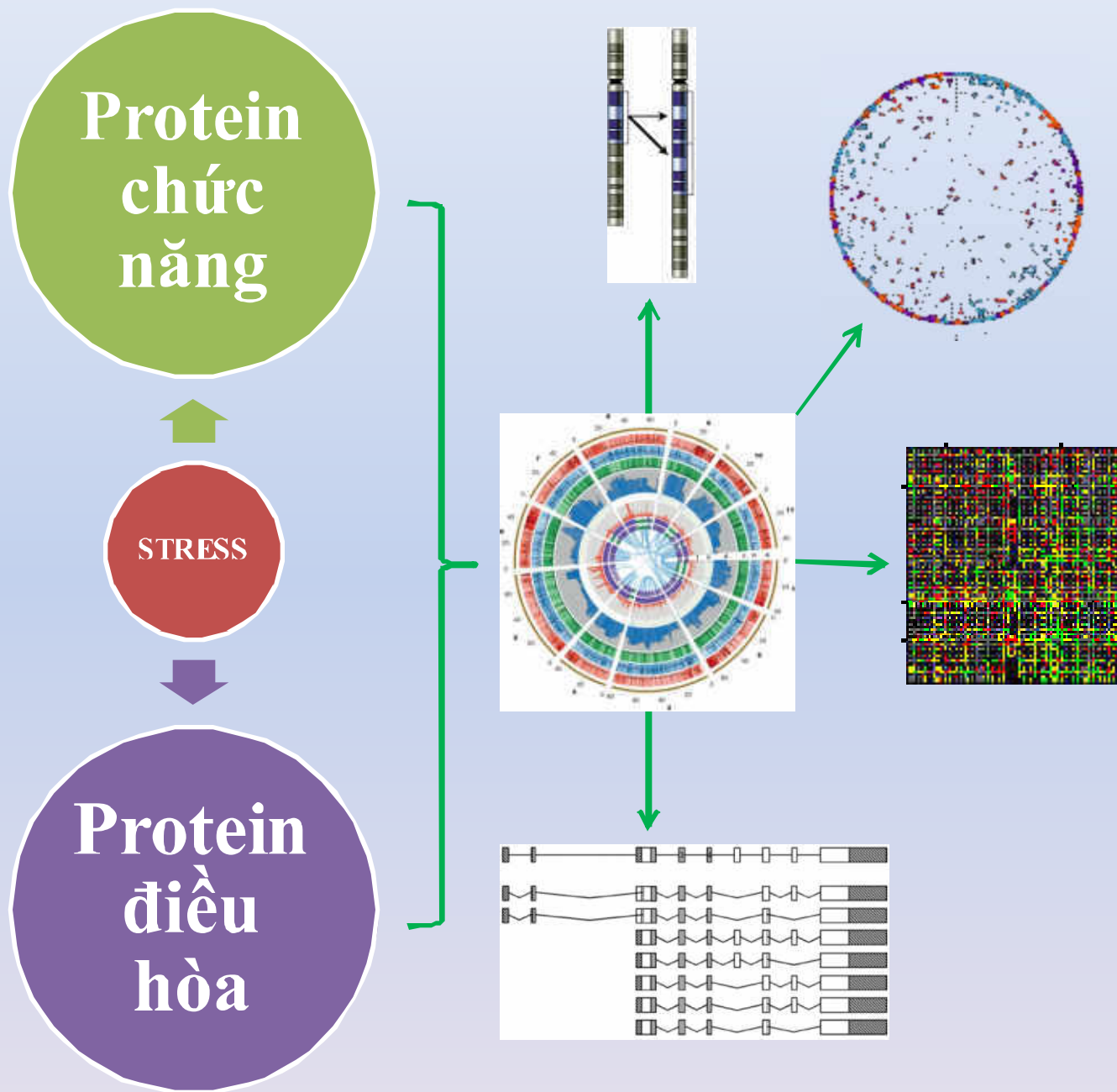
“BIO-INFORMATICS”
TIN SINH HỌC

QTL/Gene/Chỉ thị phân tử

MAS

Kỹ thuật di truyền

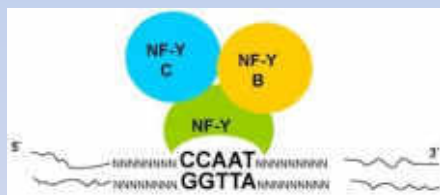
Hình 2. Sự hỗ trợ của tin sinh học trong các lĩnh vực nghiên cứu.



Hình 3. Phương pháp tiếp cận của nghiên cứu

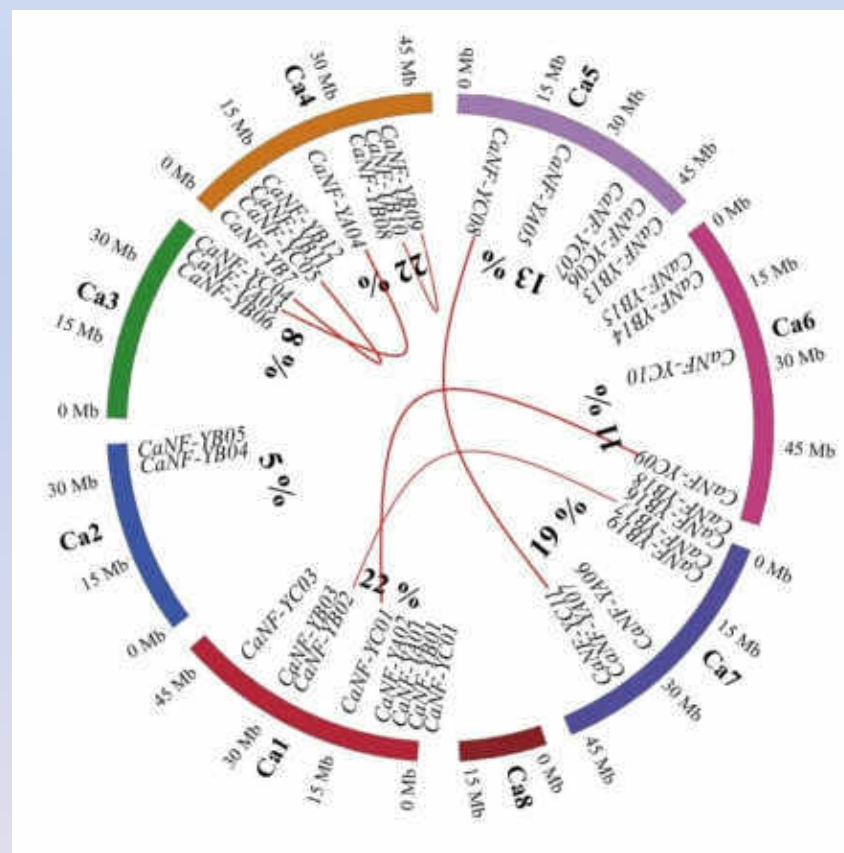
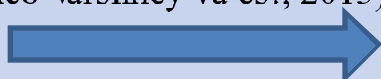
MỘT SỐ KẾT QUẢ ĐÃ ĐẠT ĐƯỢC

Xác định các họ gen mã hóa protein liên quan đến tính chống chịu ở hệ gen thực vật



Nhân tố phiên mã NF-Y,
gồm 3 tiểu phần NF-YA,
NF-YB và NF-YC

Chickpea Genome
Analysis Project
(theo Varshney và cs., 2013)



Hình 4. Họ gen mã hóa nhân tố phiên mã Nuclear factor Y trên đậu gà (*Cicer arietinum*) (theo Chu và cs., 2017)

Mô phỏng cấu trúc của enzyme MSRA



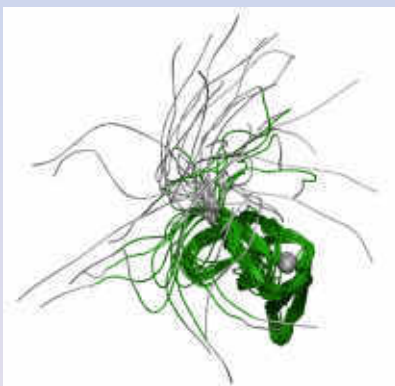
Họ Pfam: **PF01625**



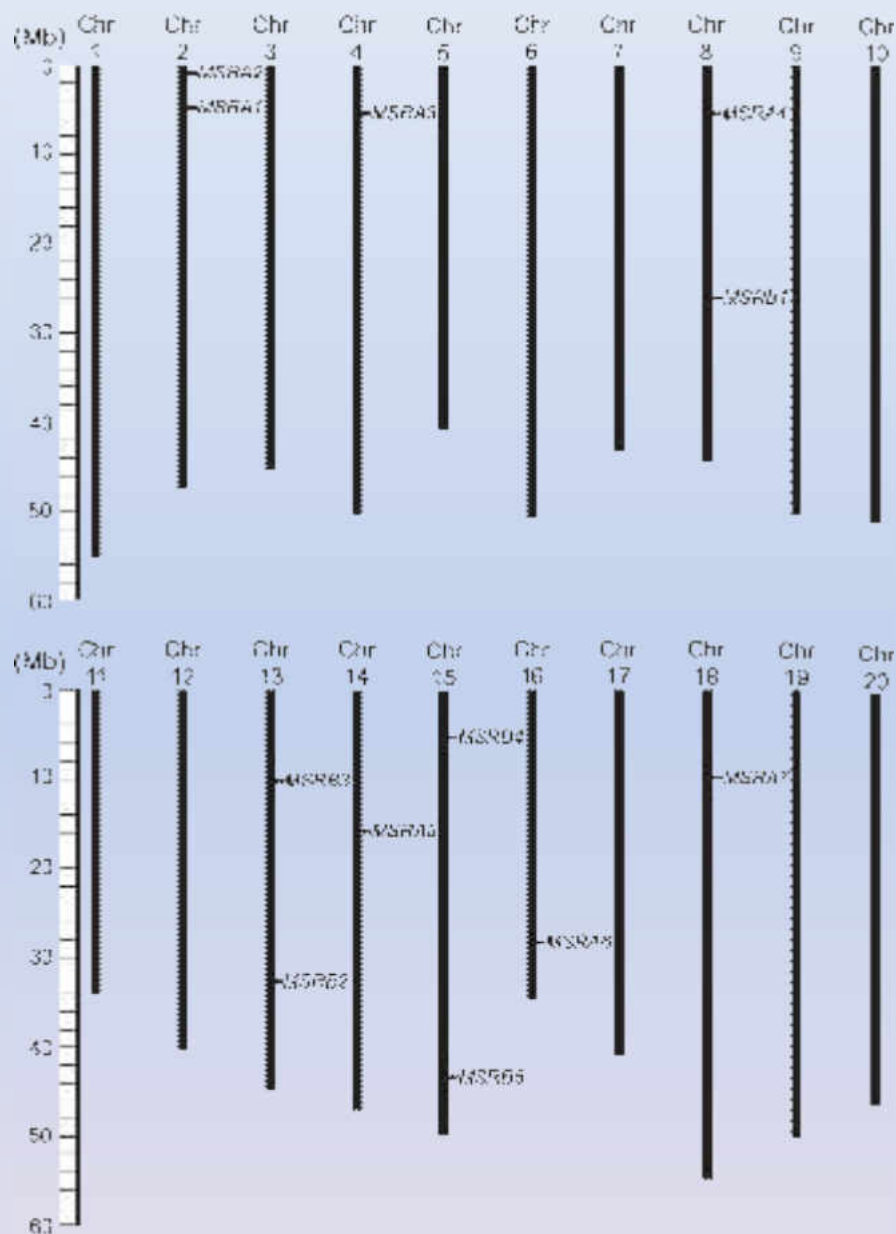
Họ Pfam: **PF01641**



Hệ gen đậu tương *Wm82.a2.v1*
(theo Schmutz và cs., 2010)



Mô phỏng cấu trúc của enzyme MSRB



Hình 5. Họ gen mã hóa enzyme methionine sulfoxide reductase trên đậu tương (*Glycine max*) (theo Lê và cs., 2013; Chu và cs., 2016)

Phân tích hiện tượng lặp gen và cơ chế nhân rộng của họ đa gen trong hệ gen thực vật

Chitinase-related agglutinin (CRA)

GeneBank: [ABL98074.1](#)

(theo Văn Damme và cs., 2016)

+

Hevein (Hvi)

GeneBank: [ABW34946.1](#)

(theo Văn Damme và cs., 2016)



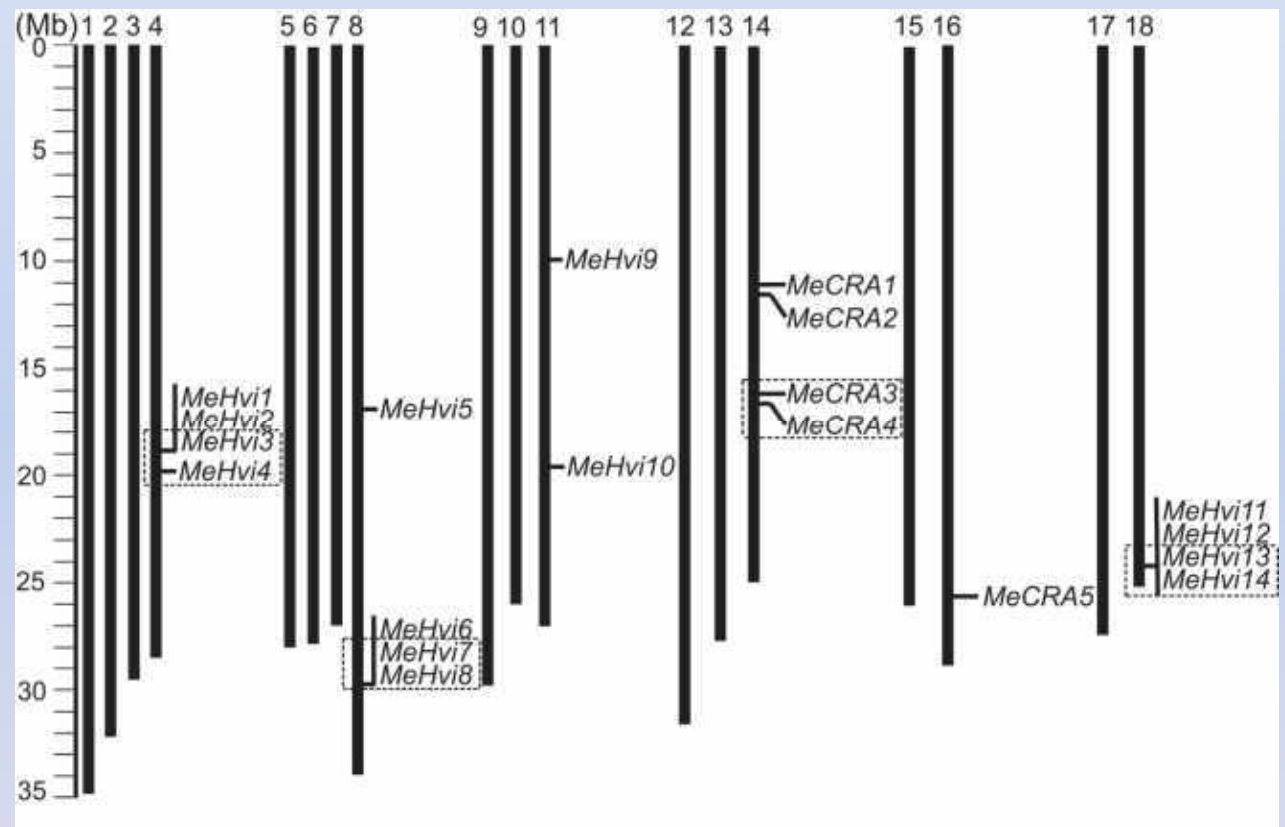
Cassava genome v6.1

Bioproject: [PRJNA234389](#)

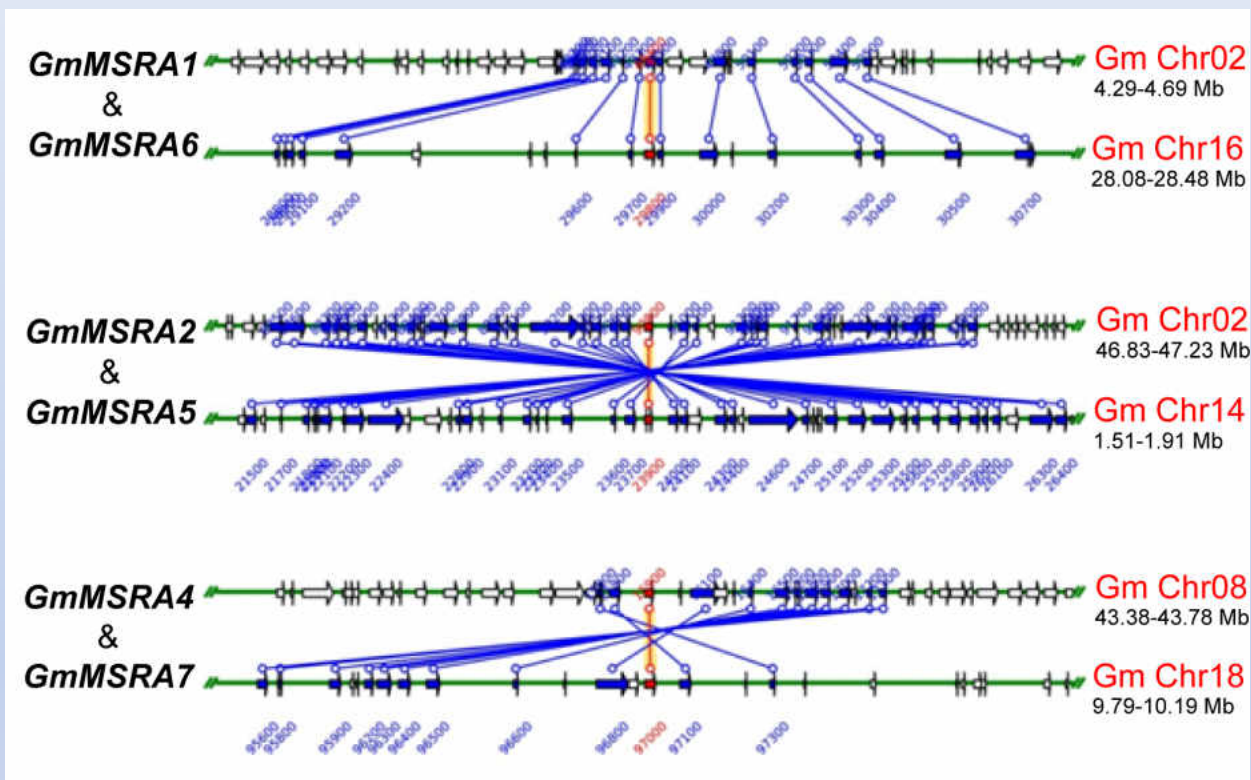
(theo Bredeson và cs., 2016)



Căn trình tự DNA
tương đồng

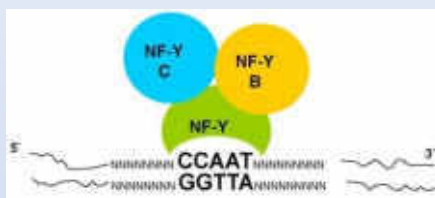


Hình 6. Họ gen mã hóa lectin trên sắn (*Manihot esculenta* Crantz) (theo Chu và cs., 2016)



Hình 7. Hiện tượng gen lặp ở họ đa gen *GmMSRA* ở hệ gen đậu tương (*Glycine max*) (theo Chu và cs., 2016, Lê và cs., 2017)

Cặp gen lặp	Mức độ tương đồng	Ks	Ka	Ka/Ks
<i>GmMSRA1/A6</i>	92,91%	0,1833	0,0319	0,1708
<i>GmMSRA2/A5</i>	96,00%	0,10422	0,0235	0,2255
<i>GmMSRA4/A7</i>	73,76%	0,0765	0,0196	,02562



Nhân tố phiên mã NF-Y,
gồm 3 tiểu phần NF-YA,
NF-YB và NF-YC)

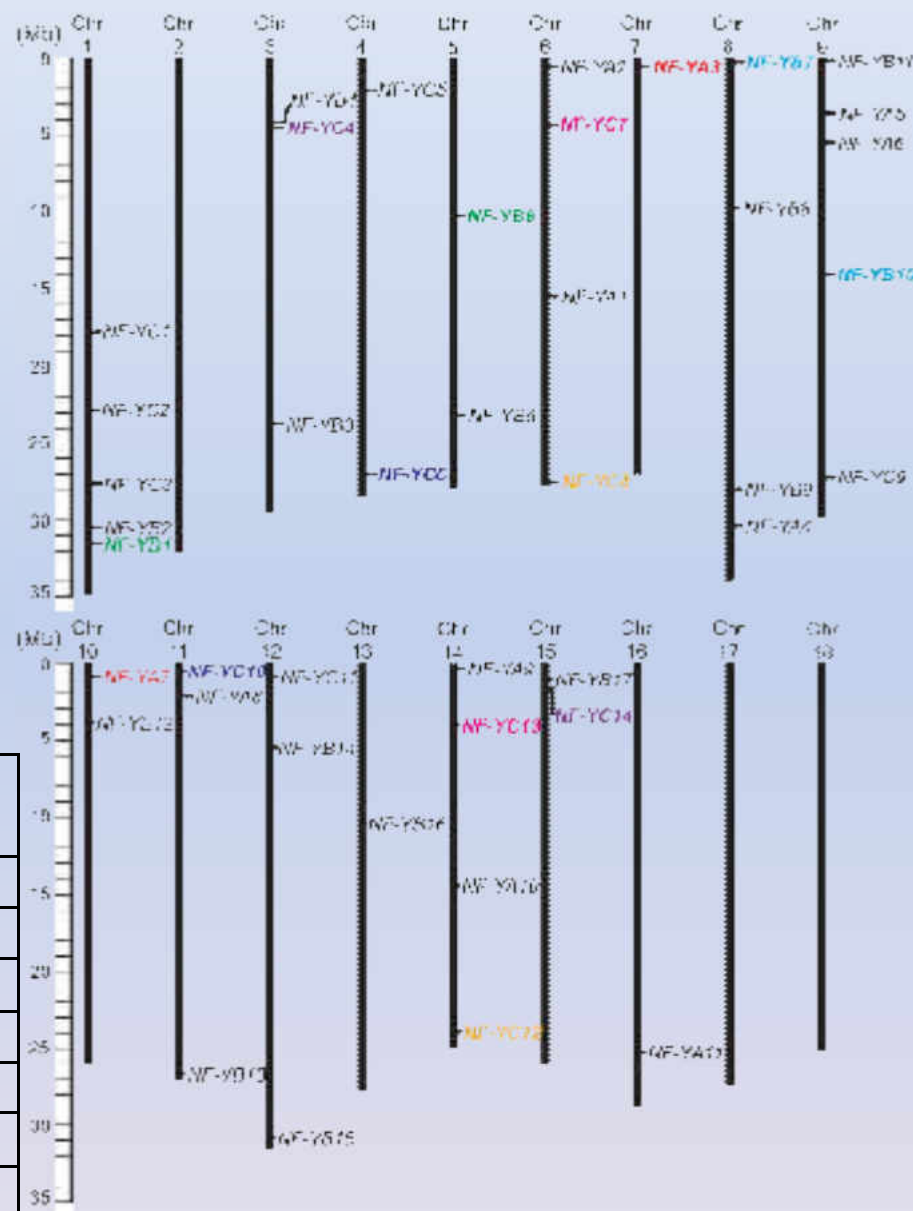


Cassava genome v6.1
Bioproject: [PRJNA234389](#)
(theo Bredeson và cs., 2016)



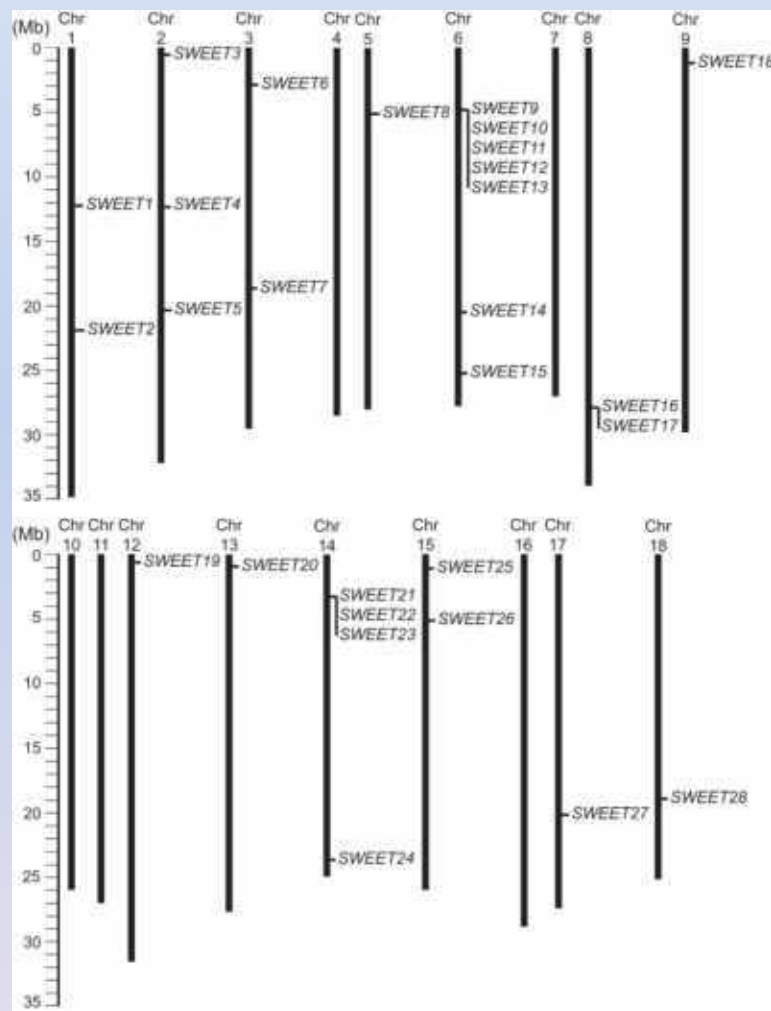
Căn trình tự DNA tương đồng

Cặp gen lặp		Mức độ tương đồng	Ks	Ka	Ka/Ks	Date (mya)
YA3	YA7	80.7	0.3150	0.1281	0.4067	24.23
YB1	YB6	80.4	0.3295	0.0192	0.0583	25.35
YB7	YB10	88.5	0.4134	0.0473	0.1144	31.80
YC4	YC14	87.3	0.3617	0.0590	0.1631	27.82
YC6	YC10	86.5	0.3430	0.0681	0.1985	26.38
YC7	YC13	91.7	0.3475	0.0221	0.0636	26.73
YC8	YC12	89.1	0.2723	0.0355	0.1304	20.95



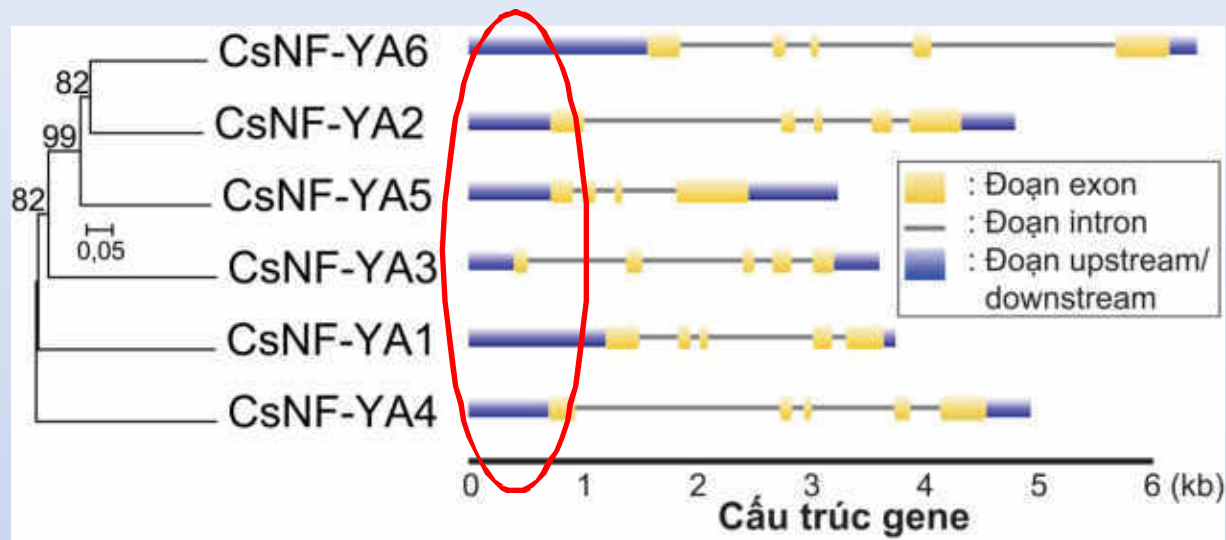
Hình 8. Hiện tượng gen lặp ở họ đa gen *MeNF-Y* ở hệ gen sắn (*Manihot esculenta* Crantz) (theo Chu và cs., 2017)

Phân tích đặc điểm và cấu trúc của họ gen mã hóa protein liên quan đến tính chống chịu ở thực vật



Hình 9. Họ gen vận chuyển sucrose *SWEET* ở sắn (*Manihot esculenta* Crantz) (theo Chu và cs., 2017)

Tên gen	Vị trí phân bố trên hệ gen	Kích thước vùng ORF	Kích thước vùng CDS	Số lượng Exon
<i>MeSWEET1</i>	Chr01R:23854836..23856816	1981	753	6
<i>MeSWEET2</i>	Chr01R:13323687..13326100	2414	783	5
<i>MeSWEET3</i>	Chr02F:702504..704395	1892	780	5
<i>MeSWEET4</i>	Chr02F:12863347..12864661	1315	798	6
<i>MeSWEET5</i>	Chr02R:13708905..13711105	2201	753	6
<i>MeSWEET6</i>	Chr03R:27968006..27970284	2279	726	6
<i>MeSWEET7</i>	Chr03R:11411434..11415892	4459	705	6
<i>MeSWEET8</i>	Chr05F:5150773..5152040	1268	816	6
<i>MeSWEET9</i>	Chr06R:23148958..23150287	1330	837	6
<i>MeSWEET10</i>	Chr06R:23142828..23144476	1649	855	6
<i>MeSWEET11</i>	Chr06R:23135496..23137211	1716	843	6
<i>MeSWEET12</i>	Chr06R:23124925..23126390	1466	849	6
<i>MeSWEET13</i>	Chr06R:23118474..23120168	1695	840	6
<i>MeSWEET14</i>	Chr06F:21384207..21385352	1146	717	6
<i>MeSWEET15</i>	Chr06F:25936409..25937653	1245	744	6
<i>MeSWEET16</i>	Chr08F:28153417..28155637	2221	759	6
<i>MeSWEET17</i>	Chr08F:28161728..28163668	1941	909	6
<i>MeSWEET18</i>	Chr09R:28230559..28232476	1918	756	6
<i>MeSWEET19</i>	Chr12F:620978..623283	2306	867	6
<i>MeSWEET20</i>	Chr13F:750620..752454	1835	858	6
<i>MeSWEET21</i>	Chr14F:3824224..3825843	1620	840	6
<i>MeSWEET22</i>	Chr14F:3834293..3836005	1713	840	6
<i>MeSWEET23</i>	Chr14F:3848619..3850071	1453	846	6
<i>MeSWEET24</i>	Chr14R:1718973..1720757	1785	750	6
<i>MeSWEET25</i>	Chr15R:20700497..20702156	1660	738	6
<i>MeSWEET26</i>	Chr15F:988957..995558	6602	714	6
<i>MeSWEET27</i>	Chr17F:20178682..20181094	2413	708	6
<i>MeSWEET28</i>	Chr18R:7456966..7459052	2087	753	6



Hình 10. Cấu trúc exon/intron của họ gen mã hóa tiểu phần NF-YA ở cam ngọt (*Citrus sinensis*) (theo Chu và cs., 2017)



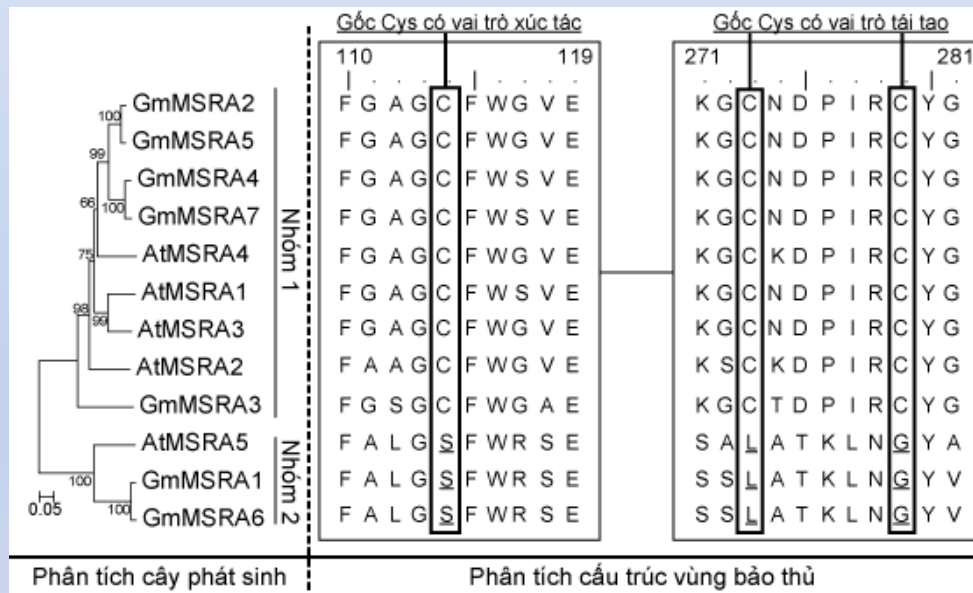
Promoter	LTRE	MBS	TC-rich repeats	HSE	ABRE	Phức hợp khởi động phiên mã	ERE	GARE	SaRE	MeJARE
<i>CsNF-YA1</i>			Y Protein điều hòa		Y	Yếu tố phiên mã chung	Y	Y		Gene X
<i>CsNF-YA2</i>			Y	Y	Y	Y	Y	Y		
<i>CsNF-YA3</i>	Y	Y	Y	Y	Y	Y	Y	Y		
<i>CsNF-YA4</i>	Y	Y	Y	Y	Y	Y	Y	Y	Y	
<i>CsNF-YA5</i>			Y							
<i>CsNF-YA6</i>			Y							

Hình 11. Khái quát cấu trúc vùng promoter của gen ở sinh vật nhân chuẩn (theo Chu và cs., 2015)

Gene name	The presences of the hormonal/stress responsive <i>cis</i> - regulatory elements								
	ABRE	GARE	AuRE	JARE	SARE	LTRE	HSE	TC-rich repeats	MBS
<i>MeCRA1</i>				+			+		
<i>MeCRA2</i>			+	+			+	+	
<i>MeCRA3</i>				+	+	+			
<i>MeCRA4</i>		+							+
<i>MeCRA5</i>		+	+				+	+	
<i>MeHvi1</i>		+	+		+			+	+
<i>MeHvi2</i>				+	+		+		
<i>MeHvi3</i>				+	+				
<i>MeHvi4</i>				+				+	
<i>MeHvi5</i>							+		
<i>MeHvi6</i>		+			+		+		+
<i>MeHvi7</i>		+	+	+			+		
<i>MeHvi8</i>	+			+		+			+
<i>MeHvi9</i>				+			+	+	+
<i>MeHvi10</i>	+			+		+	+		+
<i>MeHvi11</i>									
<i>MeHvi12</i>								+	
<i>MeHvi13</i>			+	+		+			+
<i>MeHvi14</i>			+				+		
<i>MeHvi15</i>				+			+		+
<i>MeHvi16</i>		+			+				+
<i>MeHvi17</i>	+			+					+

(theo Chu và cs., 2017)

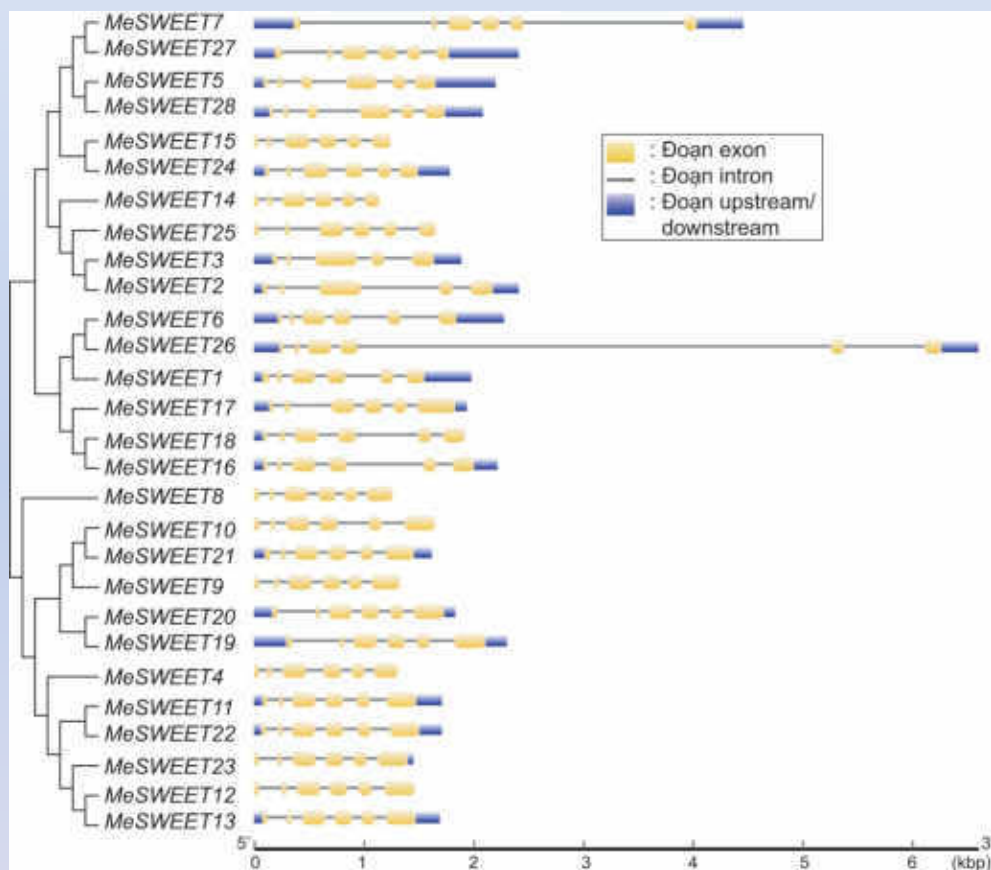
Phân tích đặc tính của các protein liên quan đến tính chống chịu ở thực vật



Tên gen	Kích thước	mW (kDa)	Pi	TargetP
GmMSRA1	250	26,17	6,14	S*
GmMSRA2	266	29,87	8,78	C*
GmMSRA3	194	21,63	5,01	
GmMSRA4	203	22,78	6,52	
GmMSRA5	266	29,69	8,74	C*
GmMSRA6	250	28,11	6,39	S*
GmMSRA7	203	22,75	5,84	M

(theo Chu và cs., 2016; Lê và cs., 2017)

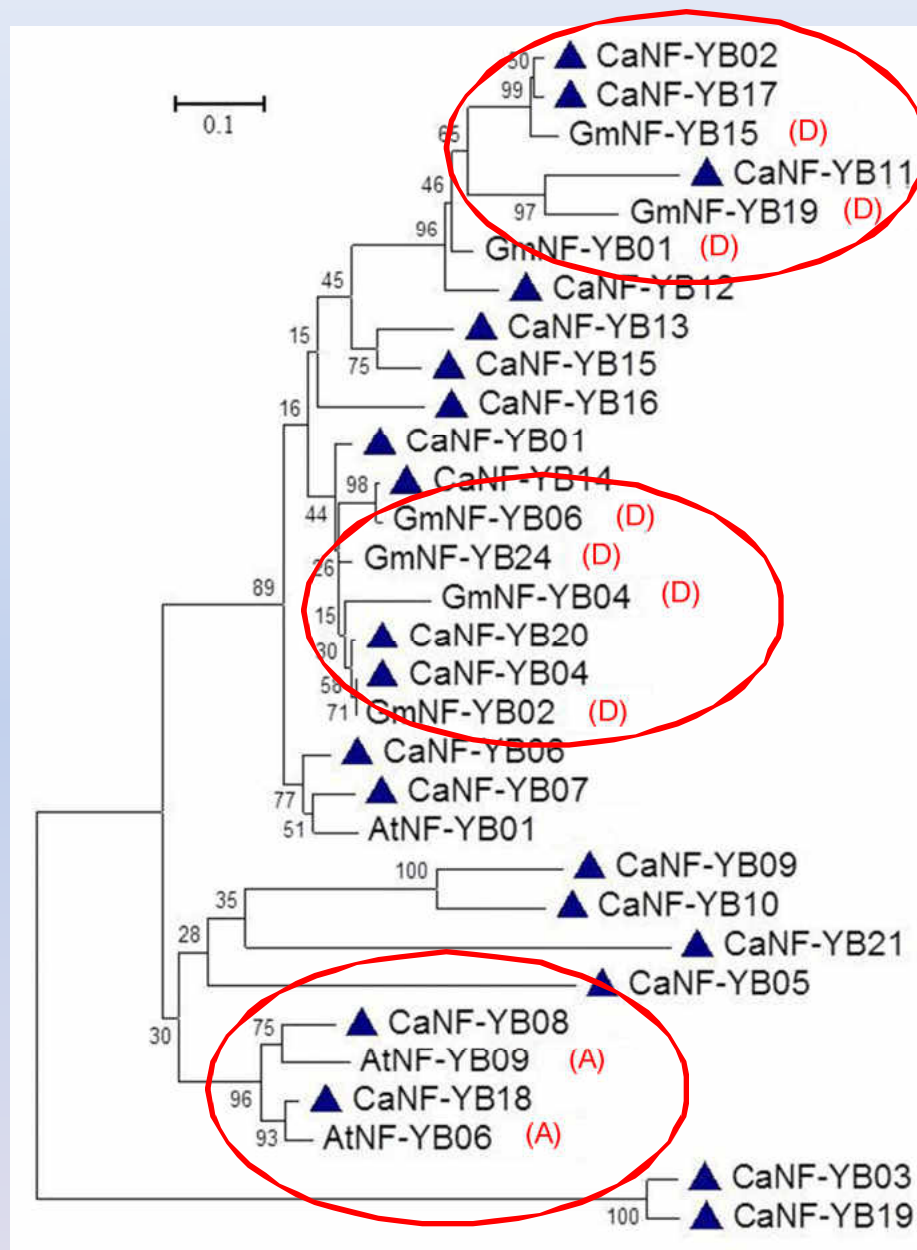
Hình 12. Sơ đồ minh họa cây phân loại và cấu trúc vùng bảo thủ cho họ gen mã hóa enzyme MSRA ở đậu tương và *Arabidopsis thaliana* (theo Chu và cs., 2016; Lê và cs., 2017)



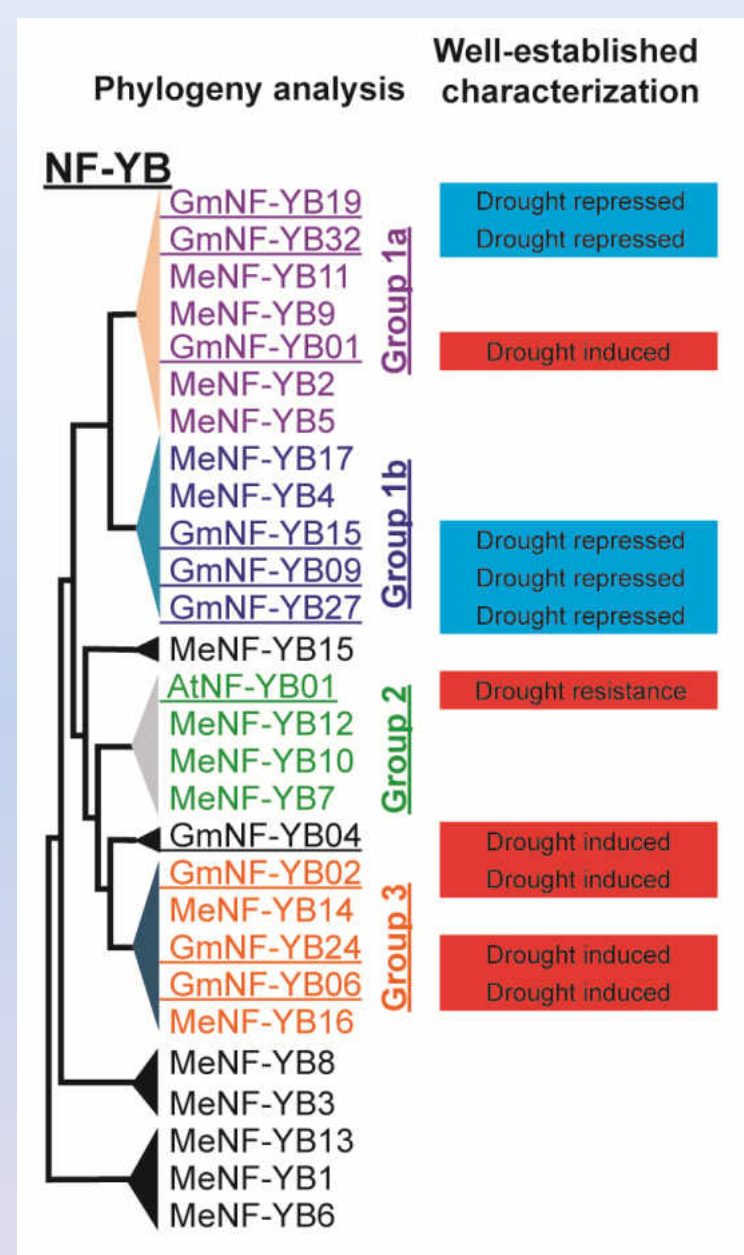
Hình 13. Họ gen vận chuyển sucrose *SWEET* ở sắn (*Manihot esculenta* Crantz) (theo Chu và cs., 2017)



Tên protein	L	pI	mW	Target P
MeSWEET1	250	8,98	27,98	S
McSWEET2	260	9,66	28,53	S
McSWEET3	259	9,58	28,49	S
MeSWEET4	265	8,62	29,77	S
MeSWEET5	250	9,30	27,68	-
MeSWEET6	241	7,72	26,33	S
MeSWEET7	234	8,63	26,07	-
MeSWEET8	271	8,43	30,28	S*
MeSWEET9	278	9,02	27,00	S
MeSWEET10	284	8,13	31,98	S*
McSWEET11	280	8,18	31,79	S*
McSWEET12	282	8,47	31,95	S*
McSWEET13	279	7,59	31,76	S*
MeSWEET14	238	8,87	31,41	-
MeSWEET15	247	9,37	27,55	-
MeSWEET16	252	9,48	27,62	S
MeSWEET17	302	9,37	33,20	S
MeSWEET18	251	9,41	27,55	S
MeSWEET19	288	6,24	32,52	S
MeSWEET20	285	7,72	32,36	S*
McSWEET21	279	8,48	31,40	S*
MeSWEET22	279	8,18	31,80	S*
MeSWEET23	281	8,91	31,71	S*
MeSWEET24	249	7,72	32,36	M
MeSWEET25	245	9,24	27,22	-
MeSWEET26	237	6,57	25,94	S*
MeSWEET27	235	8,22	25,85	-
MeSWEET28	250	9,42	27,92	S



Hình 14. Cây phân loại các tiểu phần mã hóa NF-YB ở đậu gà, đậu tương và *Arabidopsis thaliana* (theo Chu và cs., 2017)



Hình 15. Cây phân loại các tiểu phần mã hóa NF-YB ở sắn, đậu tương và *Arabidopsis thaliana* (theo Chu và cs., 2017)

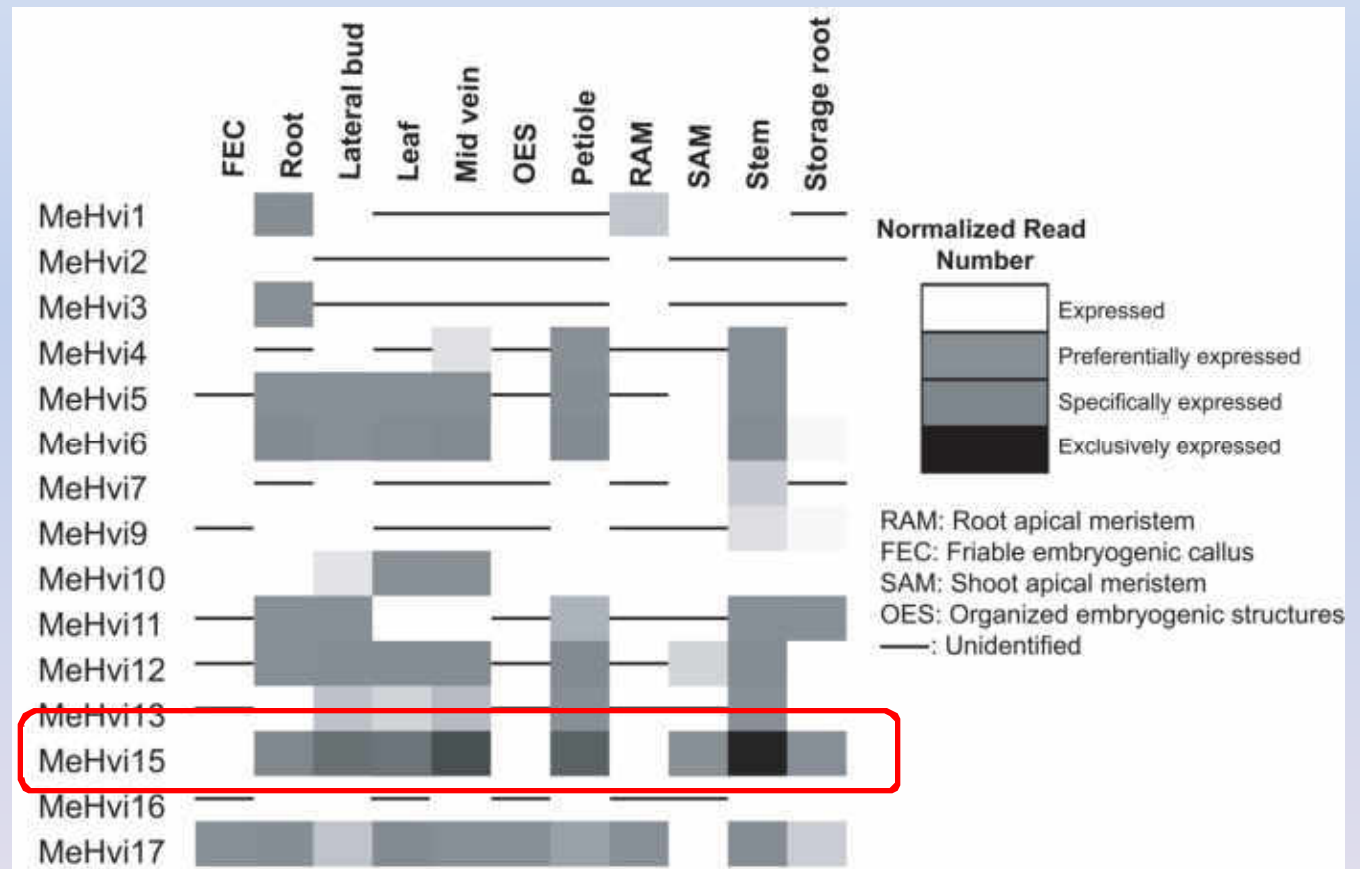
Khai thác dữ liệu biểu hiện của các gen trong mô và cơ quan chính của cây trồng

- Đánh giá mức độ biểu hiện của các gen trong điều kiện thường.

MeHvi



GEO accession: **GSE82279**



Hình 16. Dữ liệu biểu hiện của họ gen hevein trên 11 cơ quan, bộ phận chính trên cây sắn (theo Chu và cs., 2016)



Hình 17. Dữ liệu biểu hiện của họ gen vận chuyển sucrose trên cây sắn (theo Chu và cs., 2017)

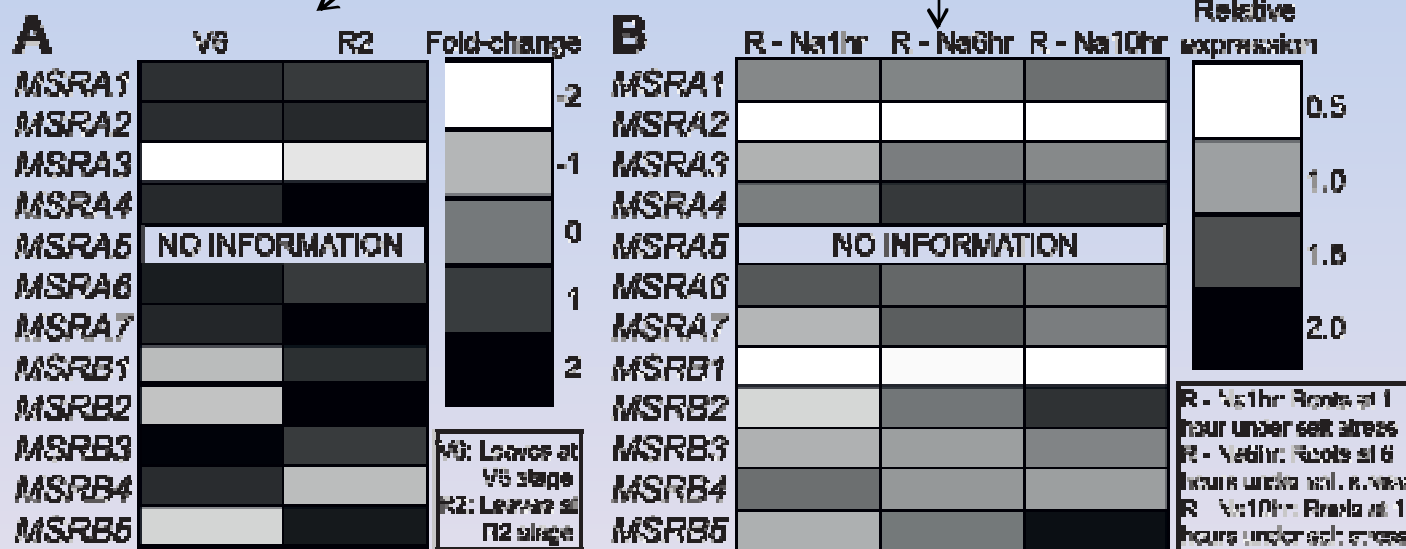
- Đánh giá mức độ biểu hiện của các gen trong điều kiện bất lợi.

GmMSRA
GmMSRB



Điều kiện hạn: (theo Lê và cs., 2012)

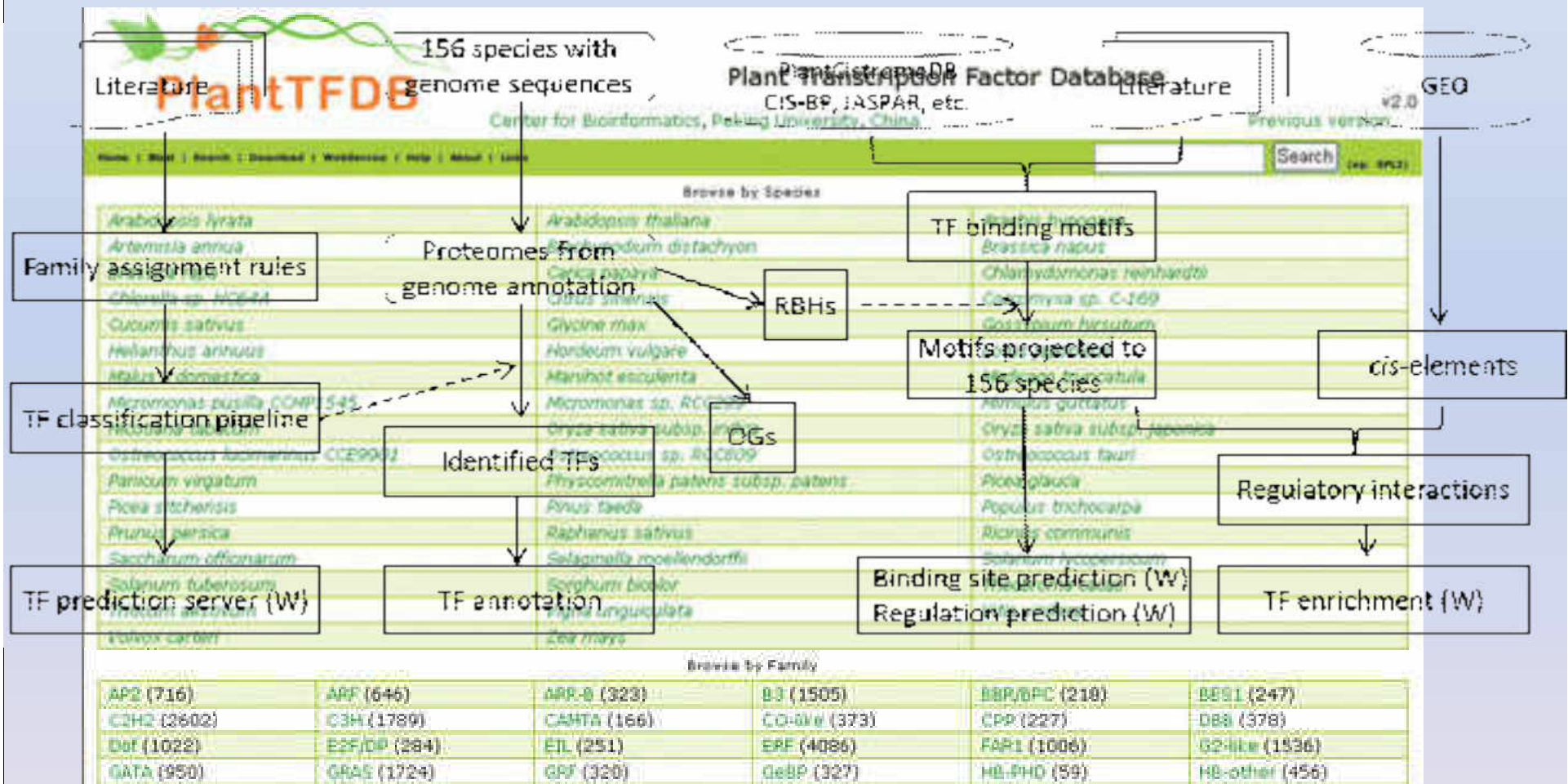
Điều kiện mặn: (theo Belamkar và cs., 2015)



Hình 18. Dữ liệu biểu hiện của họ gen hevein trên 11 cơ quan, bộ phận chính trên cây sắn (theo Chu và cs., 2016, Lê và cs., 2017)

ĐỊNH HƯỚNG NGHIÊN CỨU

Xây dựng cơ sở dữ liệu trực tuyến



Tập trung trọng điểm vào một số đối tượng cây trồng quan trọng

3K RGP *GigaScience* 2014, 3:7
<http://www.gigasciencejournal.com/content/3/1/7>

(GIGA)ⁿ
SCIENCE

DATA NOTE

Open Access

The 3,000 rice genomes project

The 3,000 rice genomes project^{1,2,3*†}

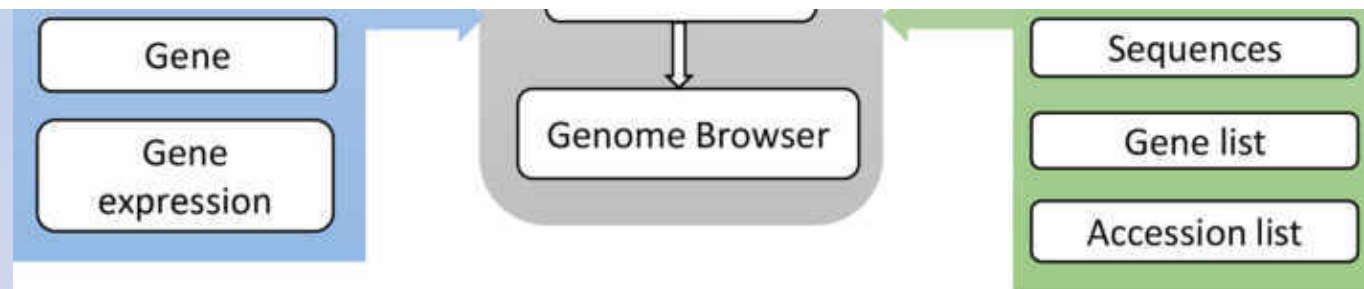
Lúa gạo (*Oryza sativa*)

- Rất gần đây...

RPAN: rice pan-genome browser for ~3000 rice genomes

Chen Sun; Zhiqiang Hu; Tianqing Zheng; Kuangchen Lu; Yue Zhao; Wensheng Wang; Jianxin Shi; Chunchao Wang; Jinyuan Lu; Dabing Zhang; ... [Show more](#)

Nucleic Acids Res (2017) 45 (2): 597-605. **DOI:** <https://doi.org/10.1093/nar/gkw958>





Jayakodi et al. BMC Plant Biology (2015) 15:138
DOI 10.1186/s12870-015-0527-0

ORIGINAL RESEARCH



RESEARCH ARTICLE

Open Access

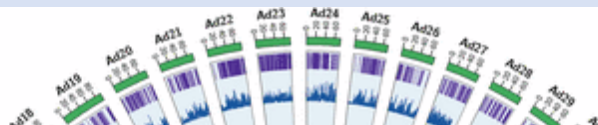


Comprehensive analysis of *Panax ginseng* root transcriptomes

Murukarthick Jayakodi^{1†}, Sang-Choon Lee^{1†}, Yun Sun Lee¹, Hyun-Seung Park¹, Nam-Hoon Kim¹, Woojong Jang¹, Hyun Oh Lee¹, Ho Jun Joh¹ and Tae-Jin Yang^{1,2*}

Amit Rai^{1*}, Mami Yamazaki^{1*}, Hiroki Takahashi², Michimi Nakamura¹,
Mareshige Kojoma³, Hideyuki Suzuki⁴ and Kazuki Saito¹

Nhân sâm (*Panax ginseng*)



Draft genome of the peanut A-genome progenitor (*Arachis duranensis*) provides insights into geocarpy, oil biosynthesis, and allergens

Xiaoping Chen^{a,1}, Hongjie Li^{b,1}, Manish K. Pandey^{c,1}, Qingli Yang^{d,e,1}, Xiyin Wang^{f,1}, Vanika Garg^c, Haifen Li^a, Xiaoyuan Chi^d, Dadakhalandar Doddamani^c, Yanbin Hong^a, Hari Upadhyaya^c, Hui Guo^f, Aamir W. Khan^c, Fanghe Zhu^a, Xiaoyan Zhang^b, Lijuan Pan^d, Gary J. Pierce^f, Guiyuan Zhou^a, Katta A. V. S. Krishnamohan^c, Mingna Chen^d, Ni Zhong^a, Gaurav Agarwal^c, Shuanzhu Li^b, Annapurna Chitikineni^c, Guo-Qiang Zhang^g, Shivali Sharma^c, Na Chen^d, Haiyan Liu^a, Pasupuleti Janila^c, Shaoxiong Li^a, Min Wang^b, Tong Wang^d, Jie Sun^d, Xingyu Li^a, Chunyan Li^b, Mian Wang^d, Lina Yu^d, Shijie Wen^a, Sube Singh^c, Zhen Yang^d, Jinming Zhao^b, Chushu Zhang^d, Yue Yu^h, Jie Bi^d, Xiaojun Zhang^e, Zhong-Jian Liu^{g,2}, Andrew H. Paterson^{f,2}, Shuping Wang^{b,2}, Xuanqiang Liang^{a,2}, Rajeev K. Varshney^{c,i,j,2}, and Shanlin Yu^{d,2}

^aCrops Research Institute, Guangdong Academy of Agricultural Sciences, South China Peanut Sub-Center of National Center of Oilseed Crops Improvement, Guangdong Key Laboratory for Crops Genetic Improvement, Guangzhou 510640, China; ^bShandong Shofine Seed Company, Jiaxiang 272400, China; ^cInternational Crops Research Institute for the Semi-Arid Tropics, Hyderabad 502324, India; ^dShandong Peanut Research Institute, Shandong Academy of Agricultural Sciences, Qingdao 266000, China; ^eCollege of Food Science and Engineering, Qingdao Agricultural University, Qingdao 266000, China; ^fPlant Genome Mapping Laboratory, University of Georgia, Athens, GA 30605; ^gShenzhen Key Laboratory for Orchid Conservation and Utilization, National Orchid Conservation Center of China and Orchid Conservation and Research Center of Shenzhen, Shenzhen 518000, China; ^hMacroGen Millennium Genomics Company, Shenzhen 518000, China; ⁱSchool of Plant Biology, University of Western Australia, Crawley, WA 6009, Australia; and ^jThe Institute of Agriculture, University of Western Australia, Crawley, WA 6009, Australia

Lạc uống biaaaaaa (*Arachis duranensis*)



Phytozome 12
THE PLANT GENOMICS RESOURCE

[JGI HOME](#) [LOG IN](#)

Species ▾Tools ▾Info ▾Download ▾Help ▾CartSubscribe

◀ Previous view

? Help with this page

Actions

↻ Revise query

🚀 Launch Jalview

🔍 Find related ... ▾

➕ Add to cart

📦 Composite family

My Data (0)

🛒 View cart

News

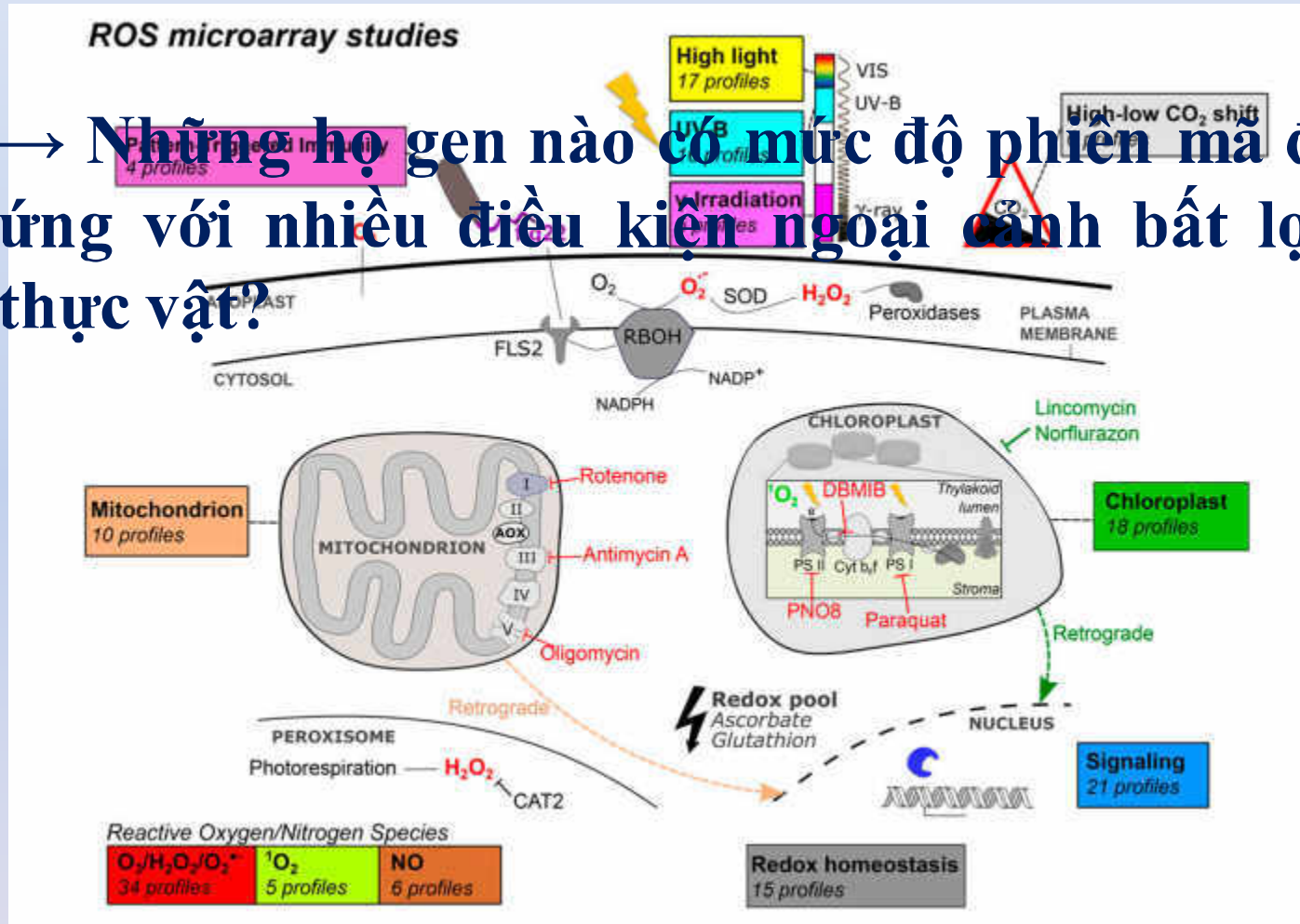
(2017-01-18) v12 Released!
Version 12 of Phytozome has been released. Find information on newly included and updated genomes in the release notes [here](#)

(2017-01-13) Coffee genome released
The genome of *Coffea arabica* from the UC Davis Coffee Genome Project is now available under Early Release and Ft. Lauderdale restrictions. View the UC Davis press release [here](#).
Bulk data download will not be available initially, but you may access individual genes and contigs via keyword searches, JBrowse, BLAST and BLAT. Please see the organism page for more information.

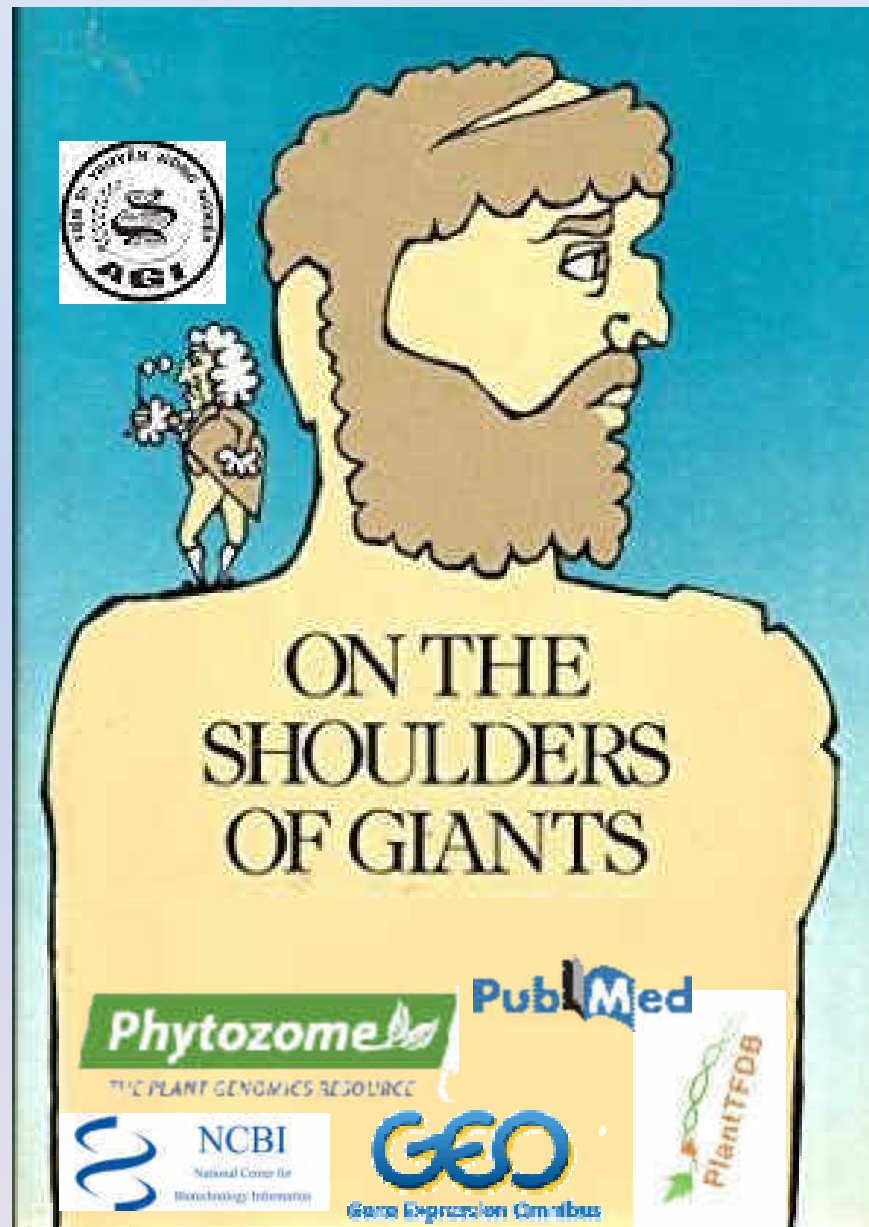
Cà phê (*Coffea arabica*)

“META-ANALYSIS”???

→ Những họ gen nào có mức độ phiên mã đáp ứng với nhiều điều kiện ngoại cảnh bất lợi ở thực vật?



KẾT LUẬN VÀ ĐỀ NGHỊ



LỜI CẢM ƠN



Prof. Lê Huy Hàm



Dr. Lê Tiến Dũng



Dr. Lam-Son Phan Tran



Dr. Phạm Thị Lý Thu



Dr. Lê Hùng Lĩnh

Viện di truyền nông nghiệp



Agricultural Genetics Institute

Chân thành cảm ơn!

Viện di truyền nông nghiệp



Agricultural Genetics Institute