

DI TRUYỀN CÂY ĐẬU NÀNH & STRESS PHI SINH HỌC

NGÂN HÀNG GEN HOA KỲ

Ngân hàng Gen của Bộ Nông Nghiệp Hoa Kỳ là nơi thu thập và bảo quản nguồn tài nguyên di truyền đậu nành lớn nhất với 18.480 mẫu giống trồng trọt, 1168 mẫu đậu nành hoang dại, được thu thập từ 84 quốc gia trên toàn thế giới.

Tập đoàn giống này đã được đánh giá kiểu gen với bộ chỉ thị phân tử SNP (SoySNP50K BeadChip) với hơn 50.000 SNPs. Tập đoàn mẫu giống vô cùng to lớn này đã được xác định và chia thành những nhóm di truyền khác nhau từ những vùng địa lý khác nhau tạo nên một tài nguyên di truyền độc đáo riêng cho lĩnh vực di truyền đậu nành. Các nhà khoa học đã phát hiện có một sự suy giảm đáng kể mức độ đa dạng di truyền trên cơ sở phân tích LD (linkage disequilibrium) và phân tích haplotype của loài đậu nành hoang dại, giống bản địa, và sự không cân đối liên kết gen của tập đoàn giống đậu nành trồng trọt vùng Bắc Mỹ. Họ đã xác định các vùng gen ứng cử viên thông qua tiến trình thuần hóa giống và tiến trình chọn lọc của con người tại Bắc Mỹ. Họ xây dựng bản đồ “block map” về haplotype trong các mẫu giống trồng trọt, hoang dại, và giống đậu nành Bắc Mỹ. Họ thấy rằng hầu hết các sự kiện tái tổ hợp xảy ra trong những vùng được đánh dấu giữa những “haplotype blocks”. Các bản đồ haplotype này vô cùng cần thiết cho nhà chọn giống để thực hiện “association mapping” nhằm mục đích xác định được những gen điều khiển các tính trạng có giá trị kinh tế quan trọng. Một trắc nghiệm chuyên môn có tên là “case-control association” làm giảm bớt các trục trặc khi xem xét vùng gen có tiềm năng trên hệ gen đậu nành, đặc biệt tại 7 nhiễm sắc thể, đó là gen điều khiển khối lượng hạt của những giống đậu nành đã được thuần hóa. Cơ sở dữ liệu này sẽ giúp cho nhà chọn giống dễ dàng hơn khi tiếp cận với ngân hàng gen, khi xác định gen điều khiển tính trạng nông học quan trọng. Cơ sở dữ liệu này giúp cho nhà chọn giống cải tiến nhanh hơn giống đậu nành cao sản trong mục tiêu nâng cao năng suất và chất lượng đậu nành (Song và ctv. 2015).

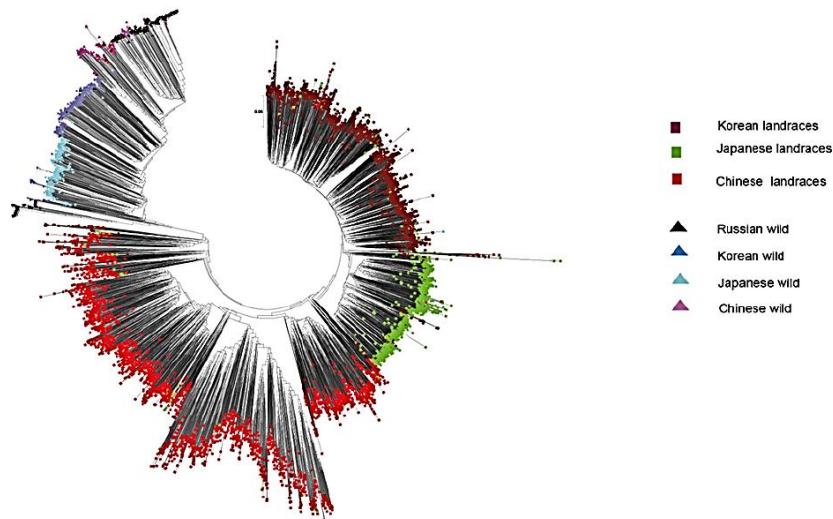


Figure 1 Dendrogram of wild and landrace genotypes from different countries.

Hình 1: Giản đồ các mẫu giống đậu nành trồng trọt và hoang dại của nhiều nước khác nhau (Song và ctv. 2015)

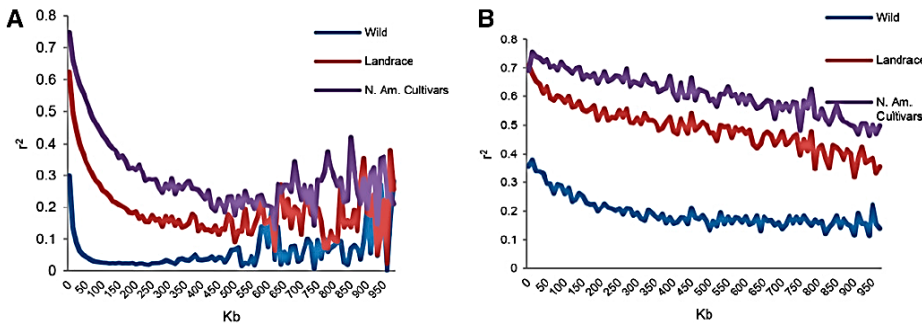


Figure 2 Linkage disequilibrium (LD) in euchromatic and heterochromatic regions. (A) LD in euchromatic regions of wild, landrace, and North American (N. Am.) cultivar soybean populations. (B) LD in heterochromatic regions of the wild, landrace, and N. Am. cultivar soybean populations.

Hình 2: Phân tích LD (sự mất đối xứng trong liên kết gen do tái tổ hợp khi chọn lọc và thuần hóa từ loài hoang dại sang loài trồng trọt) (Song và ctv. 2015)

Phân tích LD: giữa những mẫu đậu nành hoang dại gồm 806 mẫu, 5396 mẫu giống bản địa, và 562 mẫu giống đậu nành Bắc Mỹ. Chỉ có những chỉ thị SNPs với tần suất alen thấp nhất (khoảng 5%) được sử dụng để tính toán giá trị LD và sau đó hình thành những “haplotype blocks”. Tính giá trị “pairwise” của LD (r^2) trong các chỉ thị SNPs, rồi xác định “haplotype blocks” dựa trên cơ sở SNPs trong quãng có kích thước “1-Mb windows” bằng phần mềm PLINK (Purcell và ctv. 2007). “Haplotype blocks” được người ta xác định thông qua dự đoán giá trị D' cho tất cả các cặp tổ hợp của SNPs. Những cặp SNPs được xem như cùng trong một “haplotype block” nếu giá trị “one-tailed upper” đạt trên 95% mức tin cậy về thống kê và không được thấp hơn 0.7 (Gabriel và ctv. 2002).

Haplotype block là biểu thị giá trị chia sẻ có tính chất cận biên khi so sánh với các khối thuộc về loài hoang dại, loài bản địa, và loài đậu nành trồng trọt Bắc Mỹ.

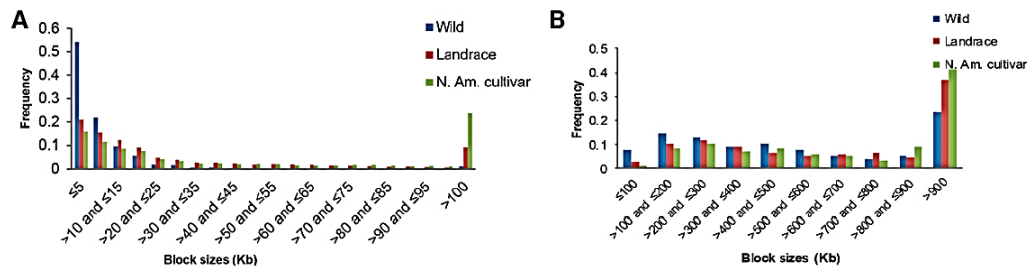
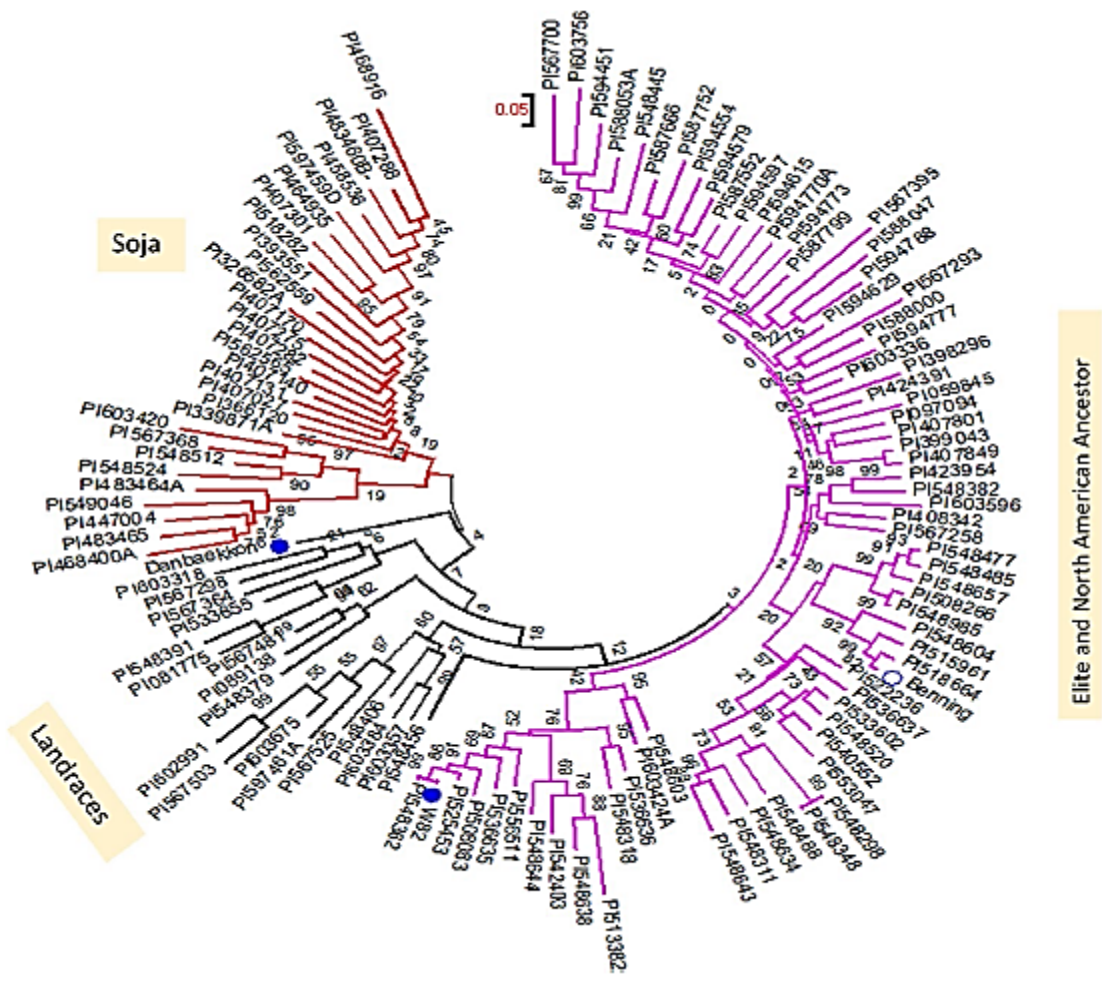
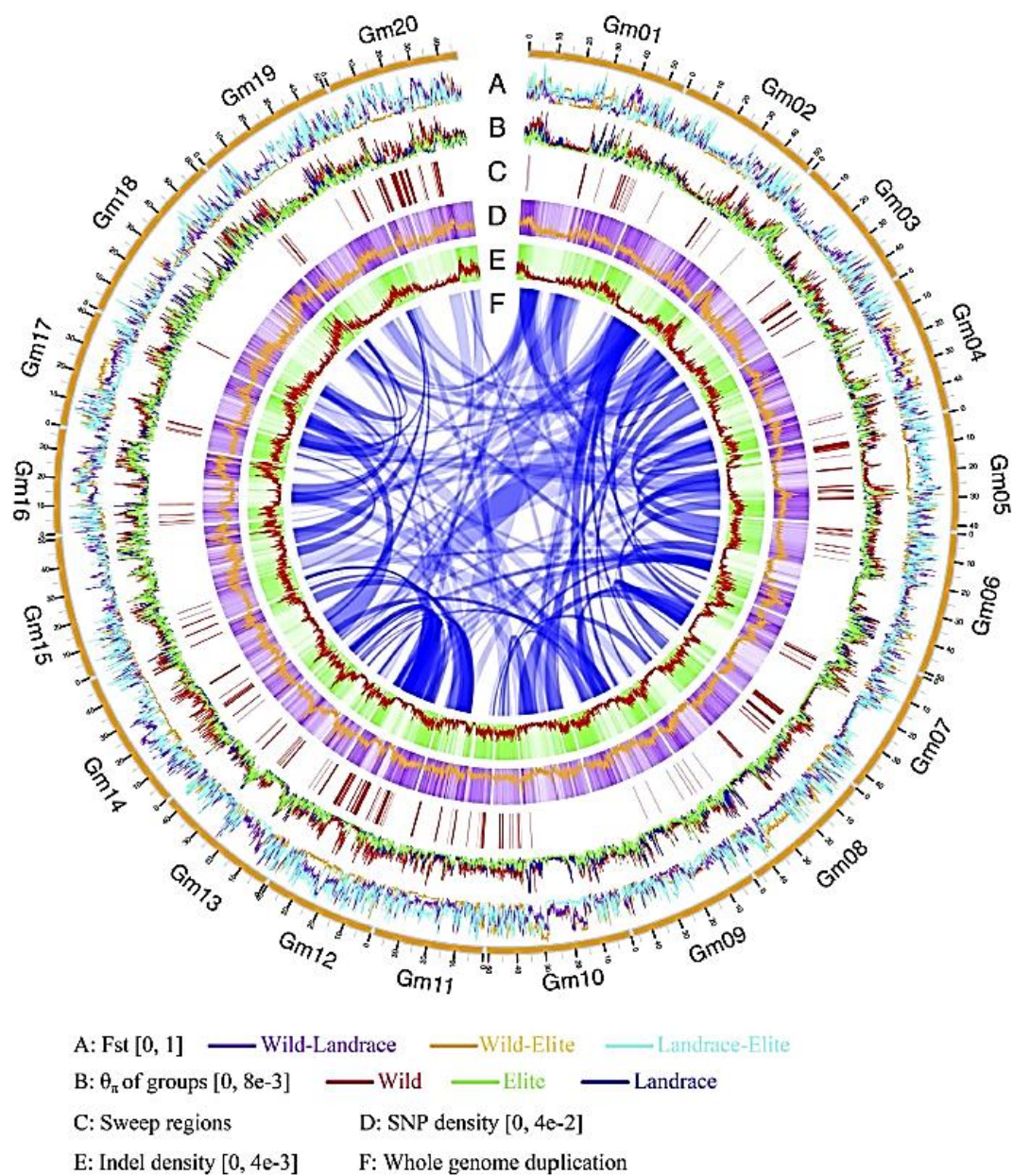


Figure 3 Distribution of haplotype block size. (A) Haplotype block size in euchromatic regions of wild, landrace, and North American (N. Am.) cultivar populations. (B) Haplotype block size in heterochromatic regions of the wild, landrace, and N. Am. cultivar populations.

Hình 3: Phân bố độ lớn của những “haplotype block” (Song và ctv. 2015).



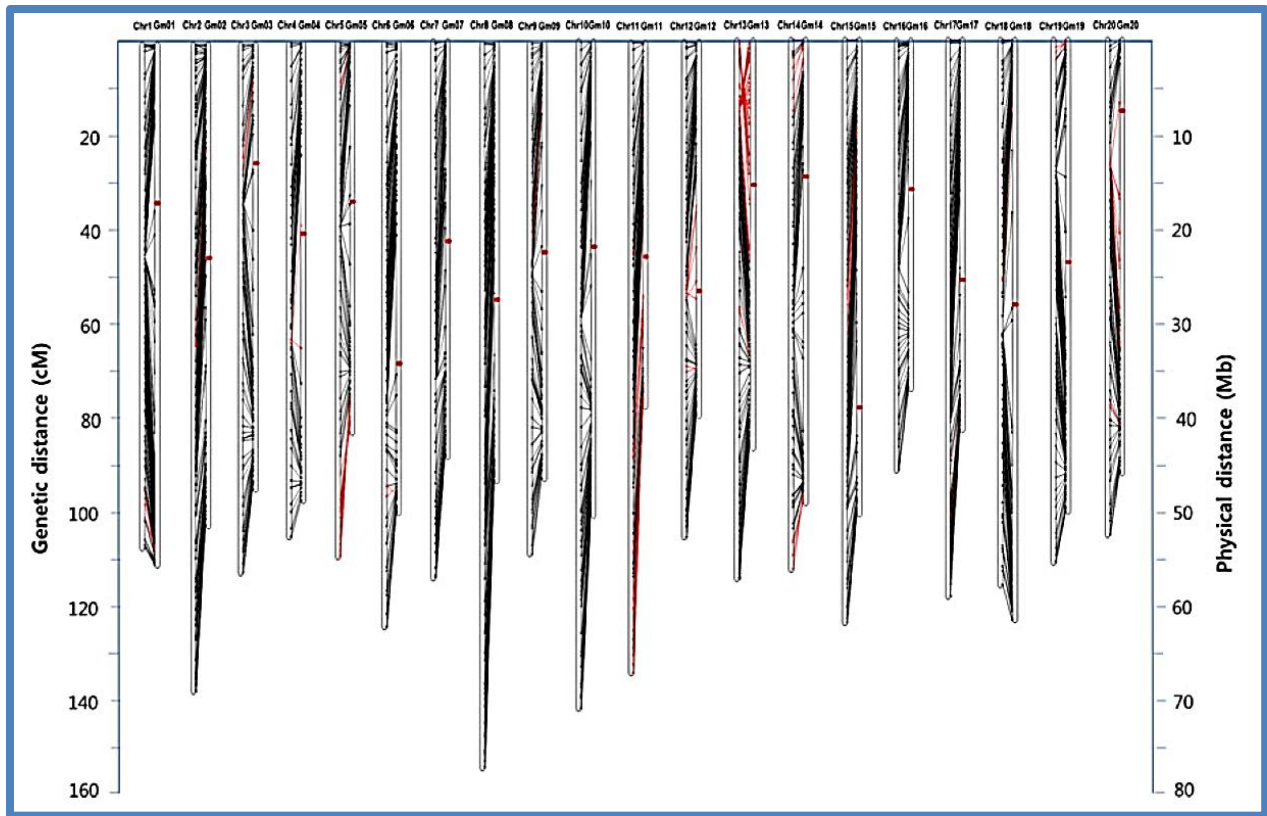
Hình 9: Phân tích SNPs trong vùng 29.8–31.6 Mb trên NST 20 cho thấy tập đoàn tổ tiên giống đậu nành Bắc Mỹ và những giống ưu việt (màu magenta) khác biệt với soja (màu đỏ) và giống bản địa của hàn Quốc (màu đen). Giống nổi tiếng của Mỹ: Williams82 và Benning, giống nổi tiếng của Hàn Quốc: Danbaekkong được ký hiệu bằng dấu chấm xanh dương (Patil và ctv. 2017).



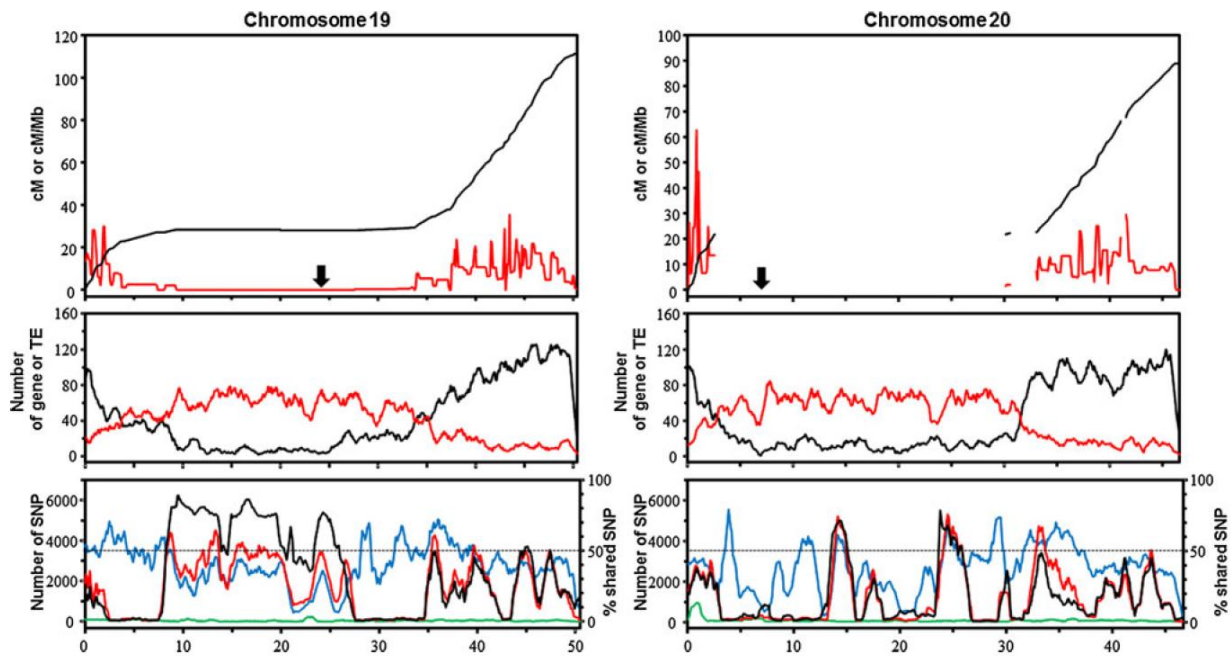
Hình 10: Cơ sở dữ liệu “resequencing” của 106 nguồn giống trong ngân hàng gen đậu nành soybean germplasm (Valliyodan và ctv. 2016).

BẢN ĐỒ VẬT LÝ TRÊN CƠ SỞ GIẢI TRÌNH TỰ DNA

Mặc dù người ta nghiên cứu khá sâu về hệ gen cây đậu nành [*Glycine max* (L.) Merrill], nhưng số cặp nhiễm sắc thể khá lớn của nó ($n=20$) giống như các loài cây trồng chủ lực khác, có thể làm ngăn cản sự phát triển bản đồ di truyền có độ phân giải cao trên cơ sở một quần thể nào đó (Lee và ctv. 2013). Do vậy, người ta đã xây dựng bản đồ trên quần thể F15 từ cặp lai giữa các dòng thuộc loài *G. max* và *G. soja* thông qua đa hình của chỉ thị phân tử InDel khi người ta tìm kiếm đa hình trong nội dung “resequencing” của hệ gen loài *G. soja*. Nhờ những chỉ thị phân tử mới InDel người ta có thể dò tìm các vùng rất nghèo chỉ thị phân tử (marker-poor regions), khoảng cách trung bình giữa 2 marker trên bản đồ giảm còn 6 cM. So sánh với trình tự tham chiếu của giống đậu nành Williams 82 điển hình, bản đồ di truyền của Lee và ctv. (2013) cho thấy thứ tự các markers thuộc 26 vùng chứa gen mục tiêu không tỏ ra nhất quán. Hơn nữa, có 7 markers định vị nhầm và hai markers vắng mặt trong giống đậu nành Williams 82, sáu markers định vị trên những “scaffolds” (khoảng trống làm khung cho nhiễm sắc thể) không hòa nhập vào những phân tử giả (pseudomolecules). Người ta xác định các chuỗi trình tự bị mất định vị tại những vị trí theo giả thiết “beginning points” tại 5 đoạn phân tử khác nhau. Những đoạn phân tử khác biệt nhau như vậy được gọi là “discordant regions” hầu hết là những sai lầm (errors) khi thực hiện kỹ thuật tập hợp lại trình tự của giống Williams 82. Sự phân bố những tái tổ hợp trên từng nhiễm sắc thể khá giống với các sinh vật khác. Đánh giá kiểu gen các InDel markers và thực hiện “resequencing” hệ gen cây đậu nành từ hai dòng bố mẹ cho thấy rằng các vùng nghèo chỉ thị phân tử có thể đặc trưng cho vùng du nhập gen mới (introgression regions). Sự phân bố markers càng dày đặc trên bản đồ di truyền sẽ đóng vai trò cầu nối giữa nghiên cứu genomics và chương trình lai tạo giống cao sản chất lượng tốt, chống chịu stress sinh học và phi sinh học (Lee và ctv. 2013).



Hình 4: Bản đồ di truyền vật lý, LK (linkage group): trái / pseudomolecule: phải (Lee và ctv. 2013)

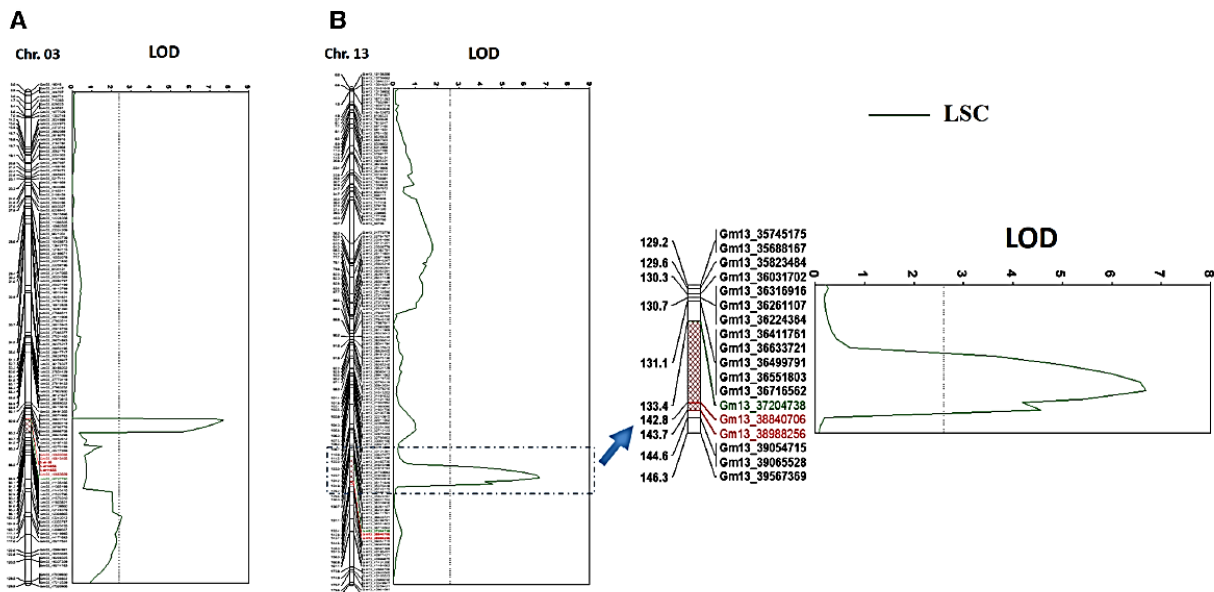


Hình 5: Đa dạng biểu thị trên nhiễm sắc thể 19–20. Trục hoành là đơn vị tính bằng triệu bp (base pairs) trên genome tham chiếu của giống Williams 82; vị trí tâm động được đề xuất bởi Schmutz

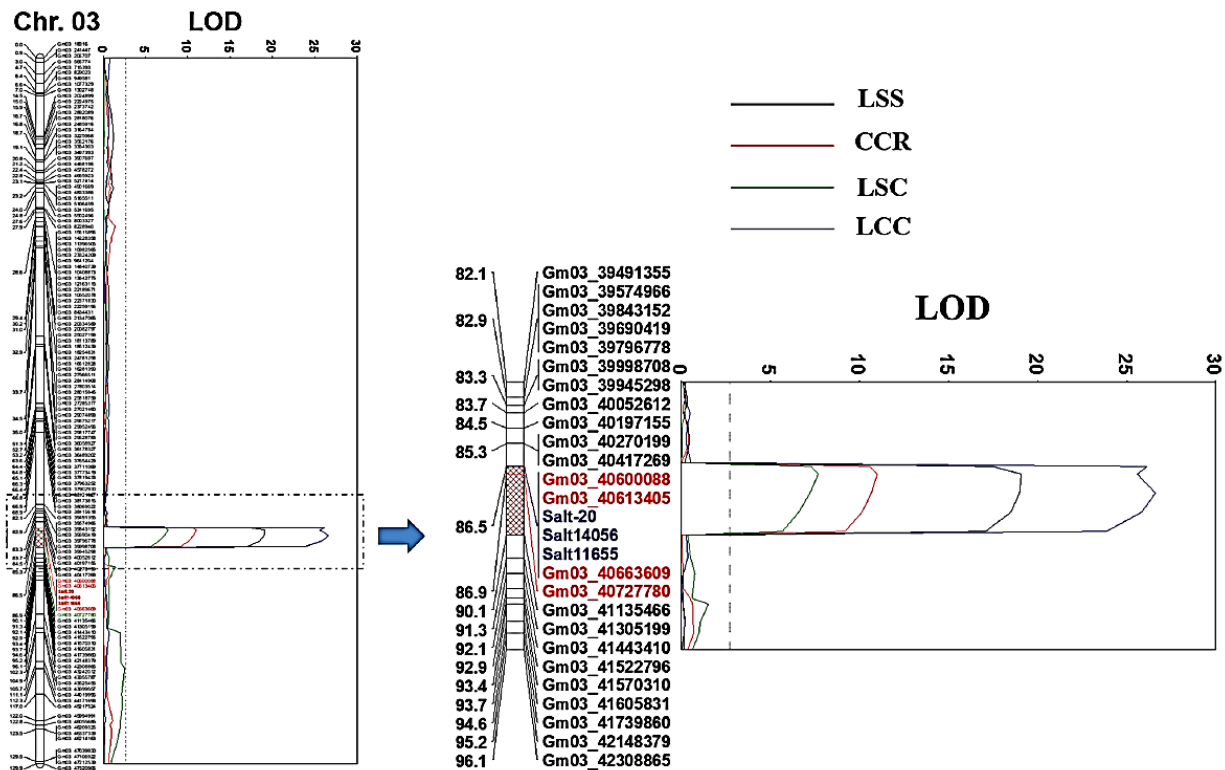
và ctv. (2010) là những mũi tên in đậm. Panel trên cùng cho thấy tương quan giữa bản đồ di truyền và bản đồ vật lý (cM, đường màu đen), mức độ tái tổ hợp tương ứng (cM/Mb, đường màu đỏ) được tính toán từ “100-kb sliding windows” đối với vùng của hệ gen giữa hai marker liền kề nhau; vùng khác biệt (discrepant regions) giữa bản đồ di truyền và bản đồ vật lý trên cơ sở trình tự DNA được thể hiện một cách không liên tục. Panel ở giữa cho thấy số gen trên từng quãng 100 kb (đường màu đen) và số nguyên tố chuyển vị TE (transposable elements) trên từng 100 kb. Panel dưới cùng cho thấy số vị trí của SNP / 100 kb (trục tung) và phần trăm SNPs được chia sẻ của giống Hwangkeum trên số SNPs đối với giống IT182932; đường màu xanh lá cây là giống are Williams 82K, đường màu xanh dương là giống IT182932, đường màu đỏ là giống Hwangkeum, và đường màu đen % là SNP chia sẻ được (Lee và ctv. 2013).

DI TRUYỀN TÍNH CHỐNG CHỊU MẶN

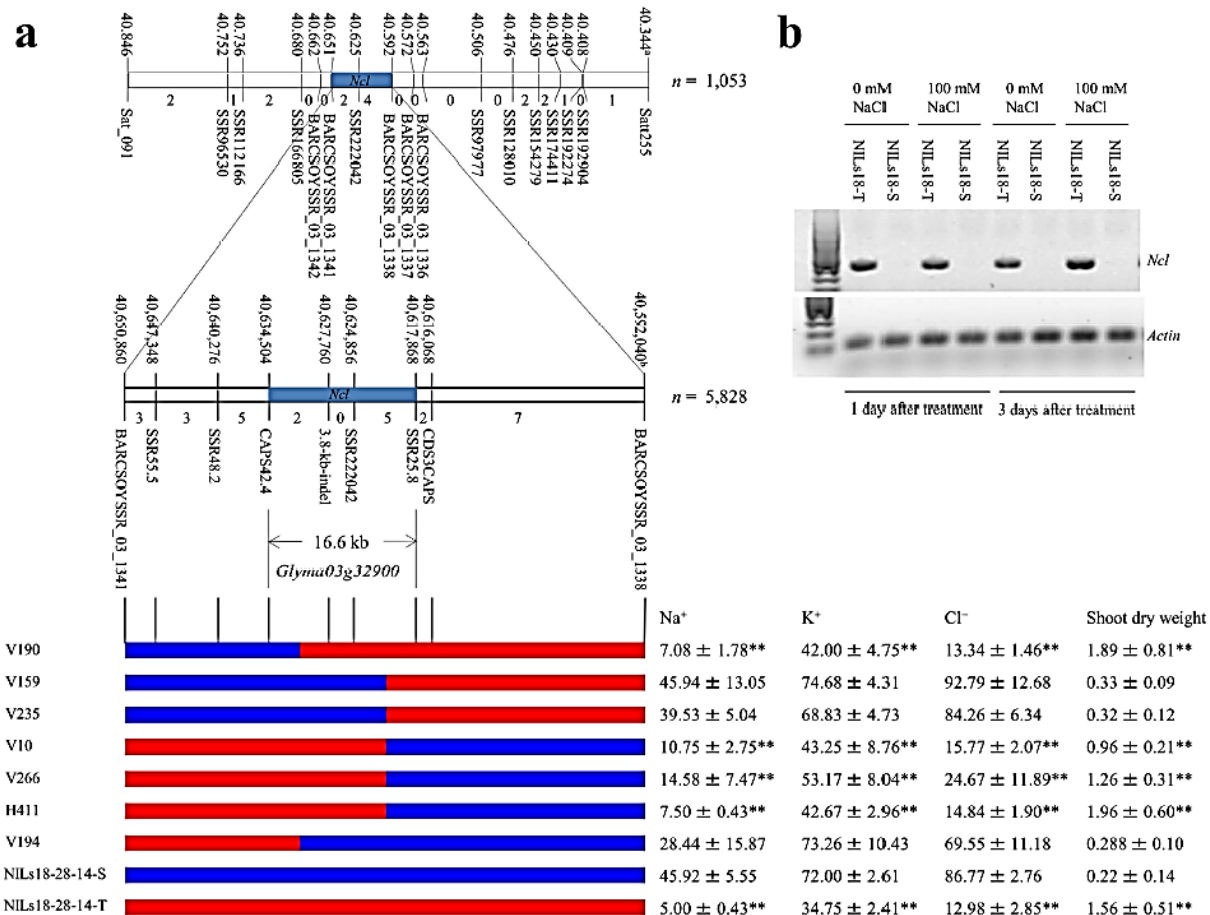
Chọn tạo giống đậu nành chống chịu mặn cao khá quan trọng cho một số vùng sản xuất đậu nành tại Hoa Kỳ và thế giới. Giống đậu nành Fiskeby III (PI 438471) thuộc nhóm có thời gian sinh trưởng 000 được ghi nhận có tính chống chịu tốt với nhiều loại hình stress phi sinh học, bao gồm mặn. Theo nghiên cứu này, quần thể lập bản đồ có 132 dòng F₂ từ tổ hợp lai của giống Williams 82 (PI 518671, nhiễm mặn trung bình) và giống Fiskeby III (chống chịu mặn) đã được phát triển để xây dựng bản đồ QTL. Đánh giá kiểu hình chống chịu mặn theo phương pháp LSS (leaf scorch score), CCR (chlorophyll content ratio), LSC (leaf sodium content) và LCC (leaf chloride content). Nghiệm thức xử lý mặn là 120 mM NaCl trong điều kiện nhà kính. Phân tích kiểu gen trên quần thể F₂ với phần mềm SoySNP6K Illumina Infinium BeadChip assay. Một alen chủ lực từ giống đậu nành Fiskeby III liên quan có ý nghĩa với LSS, CCR, LSC, và LCC trên nhiễm sắc thể 3, LOD scores là 19.1, 11.0, 7.7 và 25.6, theo thứ tự. Thêm vào đó, một locus thứ hai có liên quan đến chống chịu mặn đối với LSC đã được phát hiện và ghi trên bản đồ ở nhiễm sắc thể 13, LOD score là 4.6 và giá trị R² là 0.115. Ba chỉ thị phân tử đa hình (Salt-20, Salt14056 và Salt11655) trên NST3 biểu thị sự kết hợp mạnh mẽ với kiểu hình chống chịu mặn trong quần thể bản đồ. Những chỉ thị phân tử này sẽ rất hữu ích cho nhà chọn giống đậu nành (Đỗ Đức Tuyên và ctv. 2017).



Hình 18: LOD chỉ ra những vị trí của locus liên quan đến LSC trên NST3 (a) Một locus giả định đối với LSC trên NST 13 (b) quần thể F2: 3 họ dẫn xuất từ cặp lai giữa Williams 82 x Fiskeby III (Đỗ Đức Tuyền và ctv. 2017).



Hình 19: LOD chỉ ra những vị trí của locus liên quan đến LSS, CCR, LSC, LCC trên NST3 trong quần thể F2:3 dẫn xuất từ tổ hợp lai giữa Williams 82 x Fiskeby III (Đỗ Đức Tuyền và ctv. 2017).



Hình 20: Dòng hóa trên cơ sở bản đồ gen *Ncl* và thể hiện gen chống chịu mặn. (a) Fine mapping gen *Ncl* ở vùng có kích thước 16.6-kp giữa chỉ thị SSR25.8 và CAPS42.4 trên NST 3. (b) Thể hiện gen *Ncl* qua phân tích semi-quantitative RT-PCR trong rễ, xử lý ở ngày 1 và ngày 3 sau khi cho 100 mM và 0 mM (Control) NaCl điều kiện thủy canh. Gen actin được dùng làm đối chứng (Đỗ Đức Tuyền và ctv. 2016).

DI TRUYỀN TÍNH CHỐNG CHỊU NGẬP

Tara T. VanToai và ctv. (2001) thực hiện công trình nghiên cứu giống đậu tương chịu ngập (water logging) khi luân canh trên đất trồng lúa – như đồng bằng sông Cửu Long, do tùy triều lên xuống, do mưa nhiều, v.v...

Quần thể cận giao tái tổ hợp RILs với 208 dòng con lai từ tổ hợp lai giữa giống Archer x Minsoy và Archer x Noir I, cho ngập 2 tuần liên tục khi cây ở giai đoạn trổ bông. Đối chứng là cây không ngập nước. Địa điểm thí nghiệm tại Columbus, Ohio, năm 1997, 1998; địa điểm Wooster, Ohio năm 1998. Một QTL đã được xác định liên kết với chỉ thị phân tử Sat_064; nguồn cho gen kháng là Archer kết hợp với cải tiến tăng trưởng (11-18%) và năng suất (47-180%) trong điều kiện ngập nước. QTL có mức độ tin cậy cao ($p = 0.02 - 0.000001$) trên cả hai quần thể RILs tại Columbus (1997-1998) nhưng không tại Wooster 1998. QTL Sat_064 là độc nhất liên kết chặt chẽ với gen *Rps4* kháng đối với *Phytophthora* (*Phytophthora sojae* M.J. Kaufmann và J.W. Gerdemann). Tuy nhiên, Archer không có alen kháng *Rps4*, cho nên nó không phải là QTL

khíng bệnh. Các dòng NILs (near-isogenic lines) có và không có chỉ thị Sat_064 đã được phát triển và được thử nghiệm tính chống chịu ngập trên đồng ruộng, nhằm khẳng định sự phối hợp của QTL này với tính chống chịu ngập của cây đậu nành (Van Toai và ctv. 2001).

Table 1. Average responses of the three parents, check line and the two RI populations to flooding and control treatments at three environments.

Variety	Maturity			Early height				Late height				Yield			
	CO97	CO98	Mean	CO97	CO98	WO98	Mean	CO97	CO98	WO98	Mean	CO97	CO98	WO98	Mean
	— d after July 31 —			— cm —				— cm —				— g plot ⁻¹ —			
	Flooded														
Archer	34	41	37	26	26	25	26	39	60	56	52	95	165	126	129
Minsoy	22	32	27	22	27	25	25	25	42	45	38	5	67	103	61
Noir 1	19	26	22	28	29	28	28	35	51	54	47	11	53	70	52
Archer × Minsoy	36	40	38	21	25	25	24	31	51	55	35	42	117	98	86
Archer × Noir 1	33	39	36	24	26	25	25	36	57	58	50	42	89	113	81
IA2007	38	45	41	20	27	25	24	30	66	62	51	48	224	157	114
LSD ₀₅	2	2	6	2	2	1	1	4	5	5	4	17	45	22	19
	Control														
Archer	40	39	39	33	31	27	30	66	79	68	71	537	453	346	413
Minsoy	26	26	26	30	31	27	29	39	46	55	45	110	235	241	248
Noir 1	27	28	28	39	33	30	33	55	69	60	61	316	337	302	334
Archer × Minsoy	43	40	42	29	29	27	28	57	75	63	65	338	403	327	356
Archer × Noir 1	39	39	39	32	29	22	28	68	78	68	71	496	430	340	422
IA2007	51	47	49	27	27	27	28	66	72	79	71	616	406	458	495
LSD ₀₅	1	1	5	3	2	2	2	7	6	8	5	83	71	40	42

TABLE 1: QTL identification of waterlogging tolerance in different crops.

Species	QTL/chromosome No./genes	Marker used	Method	References
Rice (<i>Oryza sativa</i> L.)	Sub1 (Ch. 9)	SSR	MAS	[69]
Barley (<i>Hordeum vulgare</i> L.)	Qwt4-1	SSR	MAS	[70]
Barley (<i>Hordeum vulgare</i> L.)	<i>tfy2.1-1</i> , <i>tfy1.1-2</i> , <i>tfy1.2-1</i>	RFLP	MAS	[70]
Barley (<i>Hordeum vulgare</i> L.)	<i>tfy1.1-3</i> , <i>tfsur-2</i> , <i>tfsur-1</i> , <i>tfy1.1-1</i> , <i>tfmas</i> , <i>tfy2.1-2</i>	DArT	MAS	[70]
Maize (<i>Zea mays</i> L.)	Ch. 4, 9	SSR	CIM	[71]
Maize (<i>Zea mays</i> L.)	Ch. 1-3, 6, 7, and Ch. 10	SSR	CIM	[71]
Soybean [<i>Glycine max</i> (L.) Merr.]	<i>Rps</i>	SSR	SMA	[72]
Soybean [<i>Glycine max</i> (L.) Merr.]	<i>Rps</i>	SSR	CIM	[73]

*SMA: Single marker analysis, *CIM: Composite interval mapping, *MAS: Marker assisted selection.

Tổng quan của Ahmed và ctv. (2013) thể hiện trong bảng, với cây lúa nước, cao lương, lúa mạch, bắp và đậu nành.

Bhatia và ctv. (2014) liệt kê các nguồn vật liệu cung cấp gen chống chịu ngập, kiểu “waterlogging”, là JS 95 60, JS 97 52, Bhatt, Cat 3299 and JS 93 05.

Cornelious và ctv. (2005) đã phân tích bản đồ QTL tính trạng chống chịu ngập: trên hai quần thể RILs (F_{6:11}), (quần thể 1) 103 dòng thuộc A5403 × Archer; (quần thể 2) 67 dòng thuộc P9641 × Archer. Xử lý ngập nước trên đồng ruộng năm 2001, 2002, và 2003. Hệ số di truyền nghĩa rộng của tính trạng chống chịu ngập là 0.59 cho quần thể 1 và 0.43 cho quần thể 2. Sử dụng 912 chỉ thị SSRs để dò tìm các vùng ứng cử viên gen chống chịu ngập. Markers from the candidate regions were used to genotype the RILs in both populations. Cả hai phân tích SMA (single marker analysis) và CIM (composite interval mapping) đã xác định được QTL mục tiêu. Mười bảy chỉ thị trong quần thể 1 và 15 trong quần thể 2 biểu thị liên kết có ý nghĩa với QTLs ($p < 0.0001$). Trong kết quả phân tích SMA, nhiều markers liên kết với các gen *Rps* liên quan đến tính kháng *Phytophthora sojae* Kaufmann and Gerdemann. Năm markers, Satt599 trên linkage group [LG] **LG A1**, Satt160, Satt269, và Satt252 trên **LG F**, và Satt485 trên **LG N**

($p < 0.0001$). Trong kết quả phân tích CIM, một QTL tính chống chịu ngập liên kết rất gần với Satt385 trên **LG A1**, quần thể 1, năm 2003. QTL này giải thích được 10% biến thiên kiểu hình và alen này làm tăng tính chống chịu ngập từ nguồn giống Archer. Trong quần thể 2, năm 2002, một QTL chống chịu ngập định vị gần marker Satt269 trên **LG - F**. QTL này giải thích được 16% biến thiên kiểu hình và alen này làm tăng tính chống chịu ngập có nguồn gốc từ giống đậu nành Archer (Cornelius và ctv. 2005).

Các tính trạng nông học liên quan với tính chống chịu ngập được phân tích bởi Ara và ctv. (2015), tại Bangladesh [Table 1]. Ở Việt Nam, giống Nam Vang, Palmetto được xem là nguồn cung cấp gen chống chịu ngập (unpubl. data).

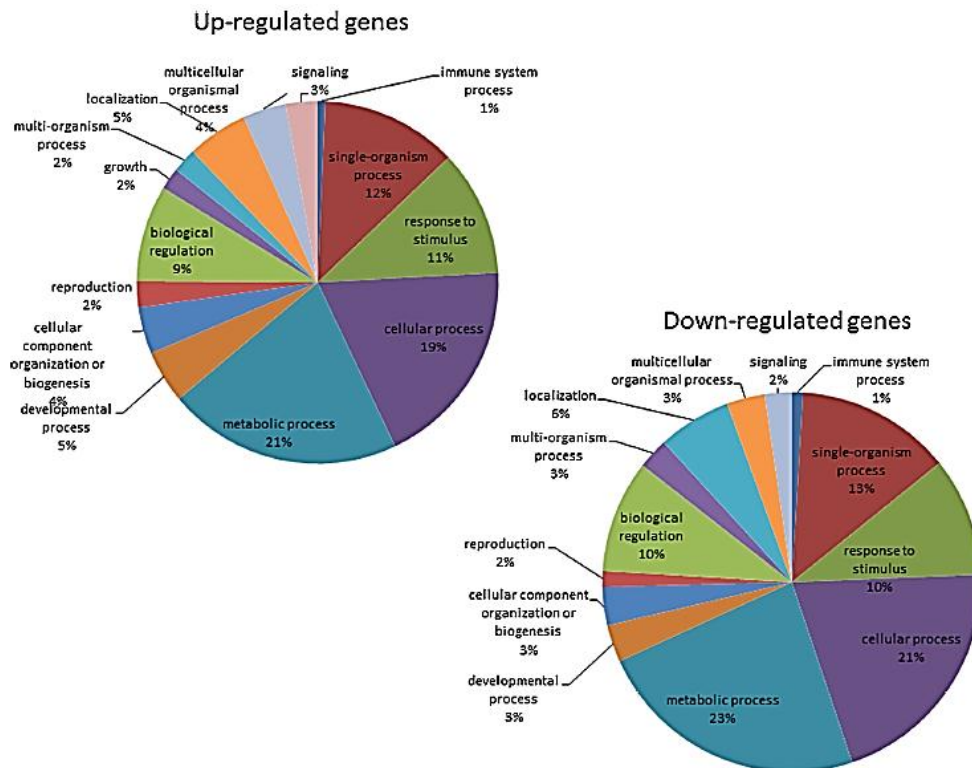
Table 1. Number of pods plant⁻¹ and number of seeds pod⁻¹ of four selected soybean genotypes as affected by waterlogging

Genotypes	Number of pods plant ⁻¹			Number of seeds pod ⁻¹		
	Control	Flooding at R1 stage	Flooding at R4 stage	Control	Flooding at R1 stage	Flooding at R4 stage
AGS 313	28.46	25.86 (90.86)	24.13 (84.78)	1.86	1.52 (81.72)	1.34 (72.04)
G 00351	20.53	17.06 (83.09)	12.93 (62.98)	1.73	1.32 (76.30)	1.15 (66.47)
BD Soybean 4	85.86	65.66 (76.47)	53.35 (62.21)	1.99	1.46 (73.33)	1.29 (64.82)
G 00197	53.26	34.86 (65.45)	25.25 (47.40)	1.89	1.22 (64.45)	1.12 (59.25)
LSD (0.05)		8.55			0.32	
CV (%)		13.20			11.42	

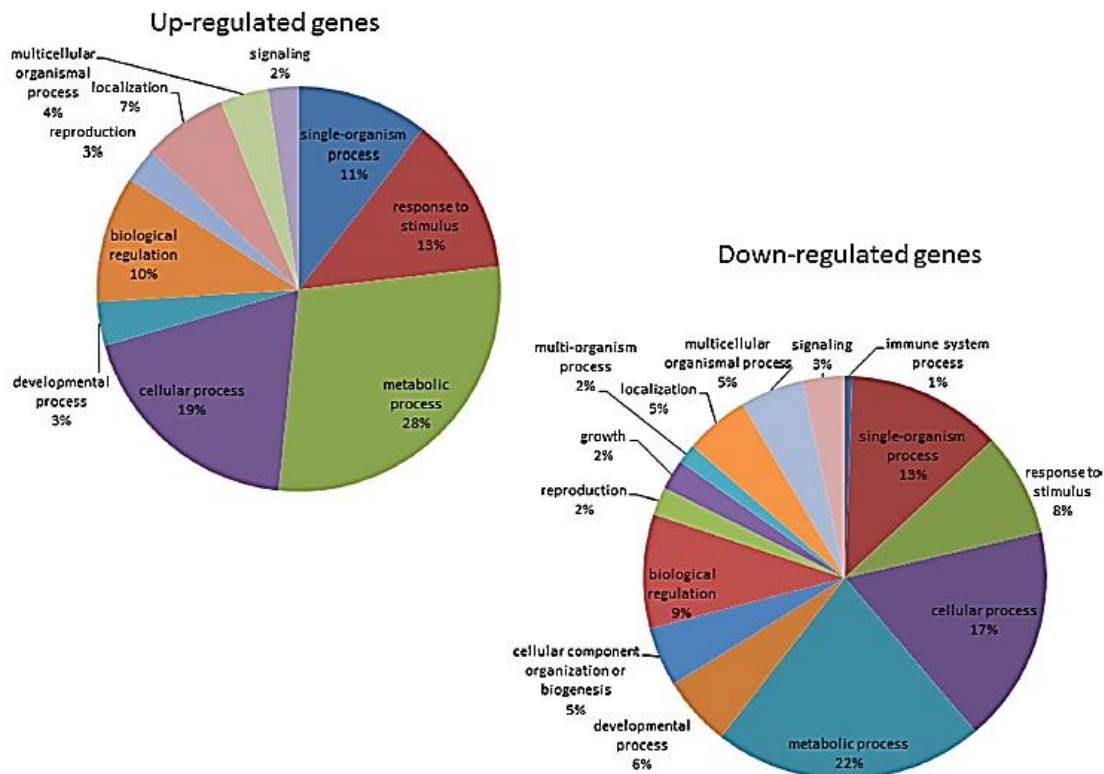
Values in parenthesis indicates per cent of control

DI TRUYỀN TÍNH CHỐNG CHỊU KHÔ HẠN

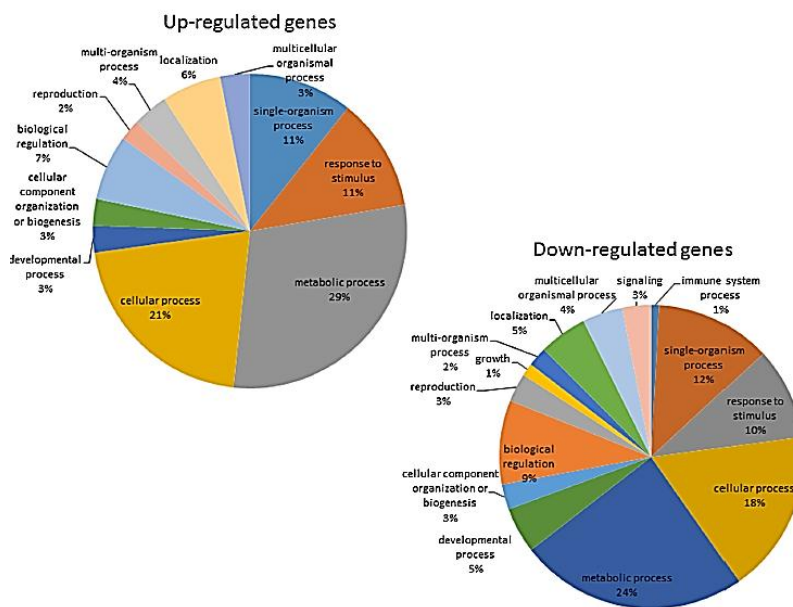
Công trình của Prince và ctv. (2015) cho thấy: Stress khô hạn gây thất thoát năng nề năng suất cây có dầu, ví dụ như đậu nành [*Glycine max* (L.) Merr]. Rất ít dòng đậu nành được xác nhận có tính chống chịu khô héo (canopy-wilting tolerant); tuy nhiên, cơ chế phân tử của chúng liên quan đến tính chống chịu này vẫn chưa được biết rõ ràng. Để hiểu được toàn bộ tiến trình sinh học như vậy, người ta tiến hành phân tích “transcriptome” của toàn hệ gen tại mô lá của hai dòng đậu nành tương nghịch nhau: dòng nhiễm khô hạn (DS) là Pana và dòng chống chịu khô hạn (DT) là PI 567690. So sánh theo kiểu “pairwise” dòng DS và dòng DT trong điều kiện xử lý khô hạn và điều kiện bình thường, người ta phát hiện ra 1914 gen và 670 gen có khả năng thay đổi gấp hai lần khi biểu hiện tính chống chịu trong nghiệm thức khô hạn. So sánh “pairwise” và phân tích độ phong phú của các gen của dòng DT cho thấy chúng điều tiết theo kiểu “down” phối hợp với kết gắn protein, kết gắn với hoạt động của hydrolase, cơ chế biến dưỡng carbohydrate / lipid, enzyme: xyloglucan endo-transglycosylases liên quan đến thành tế bào, apoplast, và protein kết gắn với diệp lục tố a/b. Mặt khác, các gen liên quan đến phản ứng stress phi sinh học, kết gắn với ion và vận chuyển ion, tiến trình ô xi hóa – khử như vậy và hoạt động vận chuyển ion được điều tiết theo kiểu “up”. Phân tích độ phong phú của gen (gene enrichment analysis) cho thấy rằng gen mã hóa UDP glucuronosyl transferase thể hiện hết sức khác biệt với nhau trong giống đậu nành PI 567690 dưới điều kiện nghiệm thức xử lý khô hạn. Các nhà nghiên cứu đã tìm thấy biến thể của chỉ thị SNPs trong họ gen “aquaporin” của dòng DT được bảo tồn trong những dòng đậu nành có mức héo rũ thấp, đây là cơ sở để thực hiện chọn giống đậu nành nhờ chỉ thị phân tử nhằm cải tiến tính chống chịu khô hạn (Prince và ctv. 2015).



Hình 2: Các tiến trình sinh học của gen DE (up and down regulated) khi bị stress khô hạn của giống đậu nành Pana (DS) RNA-seq dataset. Giá trị của mỗi chức năng sinh học thể hiện tỷ lệ gen DE trên tổng số gen đối với từng chức năng (Prince và ctv. 2015).



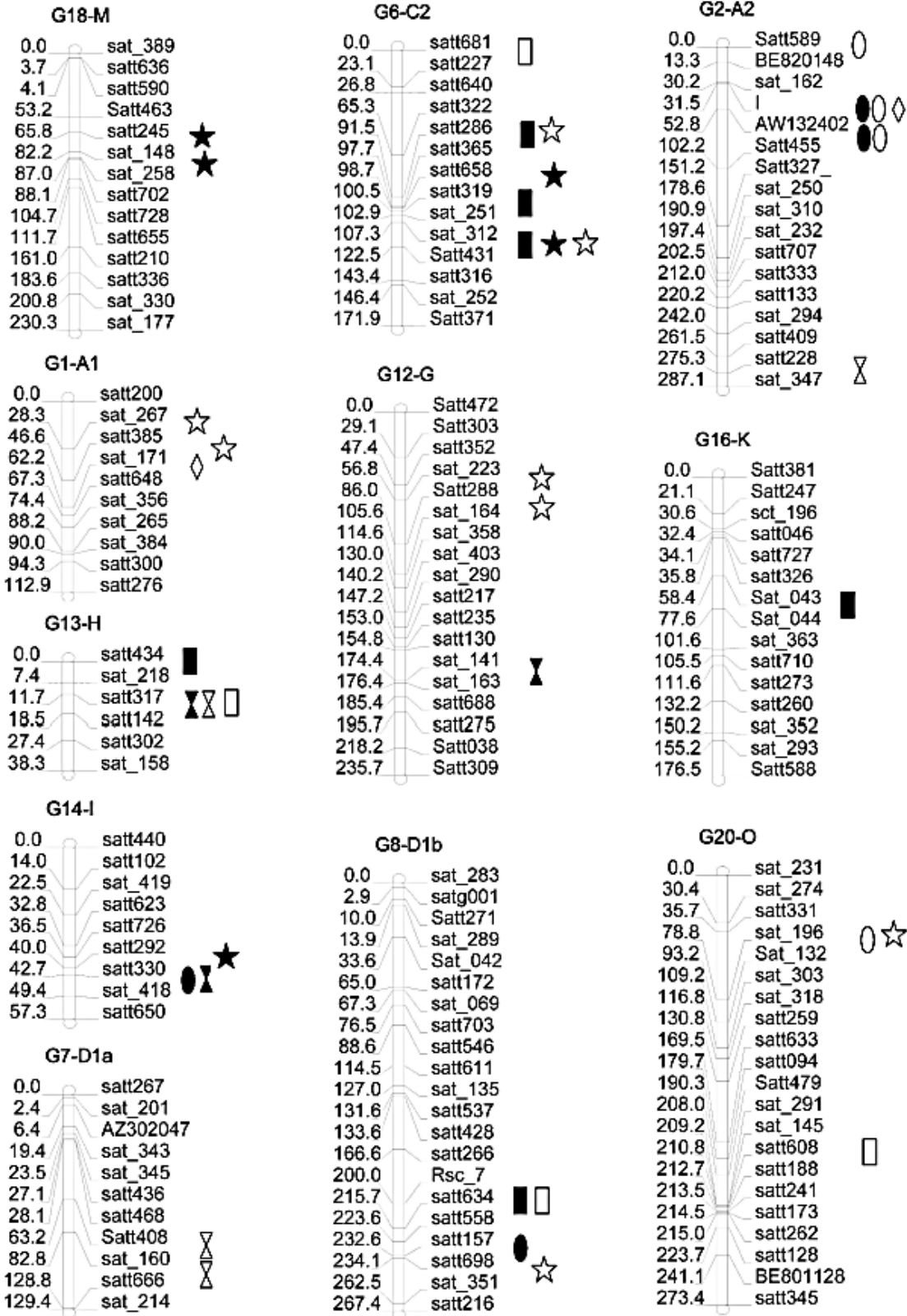
Hình 3: Các tiến trình sinh học của gen DE (up and down regulated) khi bị stress khô hạn của giống đậu nành PI567690 (DT) RNA-seq dataset. Giá trị của mỗi chức năng sinh học thể hiện tỷ lệ gen DE trên tổng số gen đối với từng chức năng (Prince và ctv. 2015).



Hình 4: Các tiến trình sinh học của gen DE (up and down regulated) khi bị stress khô hạn của giống đậu nành Pana và PI567690 (so sánh với nhau) về RNA-seq dataset. Giá trị của mỗi chức năng biểu thị tỷ lệ các gen DE trên tổng số nhóm đối với từng chức năng (Prince và ctv. 2015).

Họ thấy rằng “aquaporins” (nước và glycerol transporters), protein kết gắn với diệp lục tố a/b và TFs (FWL & YABBY) là những ứng cử viên đầy tiềm năng cải tiến tính chống chịu sự héo rũ (canopy-wilting) trong điều kiện bị khô hạn của cây đậu nành. Nghiên cứu sâu hơn cho thấy vai trò của những protein TFs (FWL & YABBY) và các thành viên của chúng tương tác với nhau sẽ cho chúng ta cơ hội tạo giống đậu nành cao sản chống chịu khô hạn. Những markers có chức năng được phát triển ở đây với chỉ thị SNPs (non-synonymous) liên quan đến aquaporins được khuyến cáo sử dụng cho MAS (Prince và ctv. 2015).

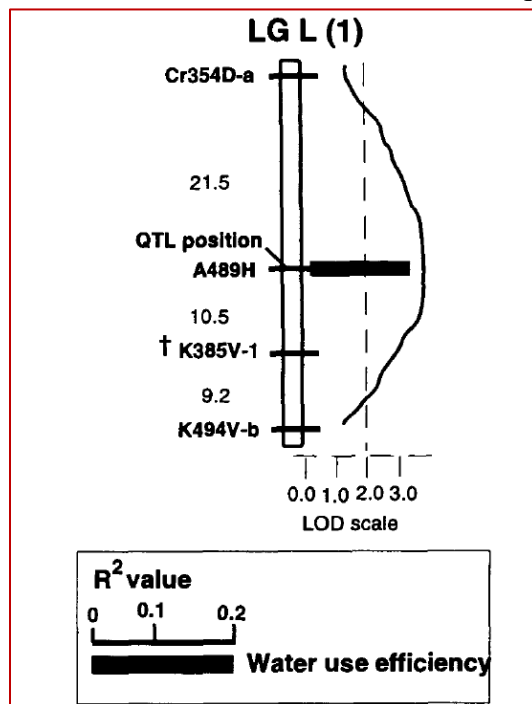
Công trình của Weijun Du và ctv. (2012) tập trung phân tích bản đồ QTL tính trạng chống chịu khô hạn trên quần thể RILs với 184 dòng F2:7:11 từ cặp lai Kefeng1 (chống chịu hạn) x Nannong1138-2 (nhễm) trong điều kiện thí nghiệm nhà lưới. Tính trạng được xem xét gồm có LWC (leaf wilting coefficient), ELWL (excised leaf water loss) và RWC (relative water content) như những chỉ thị tình trạng thiếu nước của cây đậu nành và năng suất hạt. Người ta quan sát 40 QTLs để xác nhận được 17 QTL cho tình trạng nước của lá đậu nành khi bị stress khô hạn, và 23 QTLs cho năng suất hạt ở nghiệm thức có tưới và nghiệm thức khô hạn trong nhà lưới. Hai QTLs về năng suất cho cả hai trường hợp có tưới và khô hạn trên đồng ruộng, ở nhóm liên kết gen LG - H và - D1b, trong khi 2 QTLs về năng suất nằm trên nhóm liên kết LG - C2 ở điều kiện thí nghiệm trong nhà lưới. Nhiều QTL đối với các tính trạng có liên quan đến trạng thái nước trong cây đậu nành đã được quan sát ở cả hai điều kiện thí nghiệm nhà lưới và đồng ruộng, bao gồm 2 QTL liên quan đến LWC (leaf wilting coefficient) trên nhóm liên kết LG - A2 và một QTL liên quan đến mất nước trên lá đậu nành thuộc nhóm liên kết LG - H. Tương quan kiểu hình của các tính trạng cho thấy có nhiều QTL biểu hiện gen đa tính trạng (pleiotropic) hoặc phối hợp theo kiểu liên kết giữa các vị trí (location-linked associations). Kết quả minh chứng rõ ràng cơ sở di truyền tính chống chịu khô hạn của đậu nành và cơ sở của MAS cho chương trình cải tiến giống đậu nành cao sản chịu khô hạn tốt (Weijun Du và ctv. 2012).



● → WC-F ⌘ → ELWL-F ■ → SY-F-WW ★ → SY-G-WW
 ○ → WC-G ◇ → RWC-G ⌘ → ELWL-G □ → SY-F-WS ☆ → SY-G-WS

Bảng: Bản đồ di truyền quần thể RIL - Mapmaker/EXP 3.0b. Bản đồ có 270 chỉ thị SSRs đa hình, một chỉ thị SCAR, một chỉ thị CAPS, hai chỉ thị kinh điển (*I* và *W*) và một gen kháng *Rsc-7* (kháng với SMV) tại 25 MLGs với tổng chiều dài được bao phủ trên toàn bộ genome là 3312 cM (Weijun Du và ctv. 2012).

Công trình của Mian và ctv. (1998) tập trung phân tích hiệu quả sử dụng nước của cây đậu nành (Water Use Efficiency: WUE). Trước đó, họ đã phân lập được 4 QTLs trong quần thể con lai F₄ của cặp lai giữa 'Young' x PI416937. Sau đó, họ thực hiện trên quần thể F₂ nhằm tìm hiệu tính trạng WUE QTL từ cặp lai giữa 'S100' x 'Tokyo' và xác định bản chất của WUE QTL thông qua hai quần thể con lai nói trên. Sử dụng chỉ thị RFLP để xây dựng bản đồ di truyền với 142 markers và 116 “Frderived lines” của quần thể con lai. Bản đồ phủ trên 1100 cM của genome



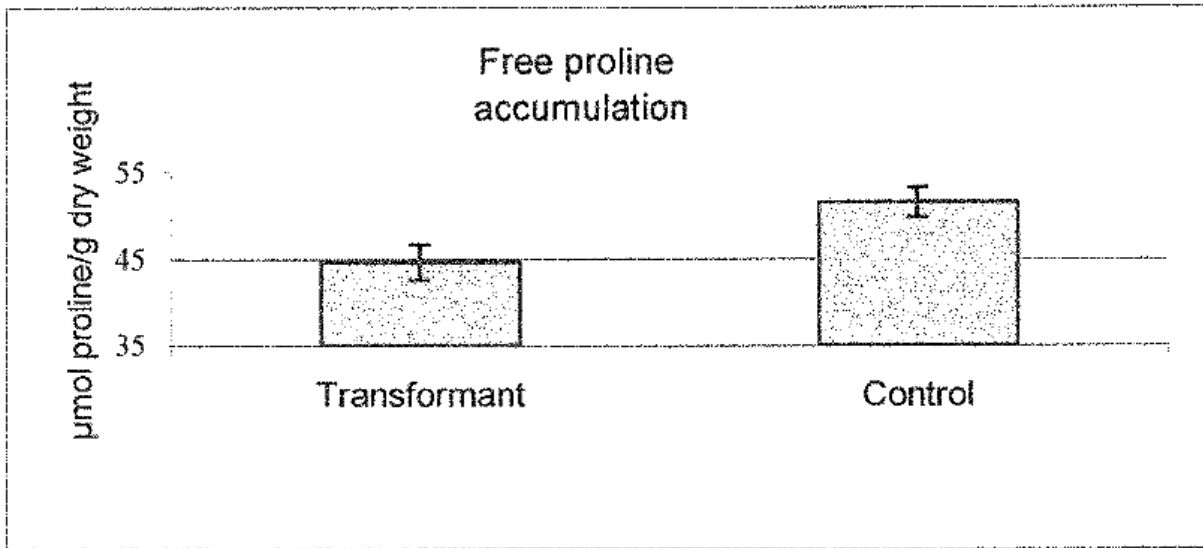
cây đậu nành với 120 markers liên kết trong 25 LGs (linkage groups). Hai mươi hai markers không có liên kết nào. Tính trạng WUE của bố mẹ và dòng con lai F₂ được đánh giá trong nhà lưới tại Athens, GA, năm 1996. Giống đậu nành S100 cho ra 9% chất khô / đơn vị nước sử dụng tại Tokyo. Các dòng con lai có giá trị WUE cao hơn S100 đã được xác định trong quần thể con lai. Hai markers độc lập kết hợp với tình trạng WUE. Không có tương tác giữa hai markers. Một (A063E) của hai markers độc lập kết hợp với tính trạng WUE trong quần thể con lai Young x P1416937. Marker A489H, trong LG - L, là chỉ thị độc nhất đối với quần thể S100 x Tokyo, giải thích được 14% biến thiên kiểu hình tính trạng WUE. kết quả phân tích bản đồ cách quãng cho thấy với MAPMAKER/QTL là vị trí QTL khả dụng nhất nằm trên LG - L tại locus A489H. Giá trị LOD (likelihood of odds score) đối với sự có mặt của QTL ở vị trí này là 3.4. Do đó, các nhà nghiên cứu này đã xác định được QTL đối với tính trạng WUE trên LG - L (Mian

và ctv. 1998).

Hình 2: RFLP linkage group L (LG - L) của quần thể S100 × Tokyo (Mian và ctv. 1998)

Công trình của J.A. de Ronde và ctv. (2000): nghiên cứu sự tích lũy của proline khi bị stress khô hạn. Proline trong cây được sinh tổng hợp từ glutamic acid và ornithine, nhóm tác giả đã sáng tạo ra dây antisense của hệ gen cây đậu nành với một gen **L-11-pyrroline-5-carboxylate reductase** (ký hiệu là **P5CR**). Sự thể hiện gen này cho kết quả hàm lượng proline được sản sinh ra, với việc sử dụng “heat shock promoter” (IHSP) rất nhạy cảm. Hoạt động của IHSP làm bất hoạt gen **P5CR**, làm cho sinh tổng hợp protein bị giảm theo. Cây đậu nành “antisense” cho chúng ta kiến thức mới về tương quan giữa hàm lượng proline, stress khô hạn và stress áp suất thẩm thấu. Stress do mannitol ở 32 and 42 °C làm tăng cường sự tích lũy proline trong cây đối chứng, trái lại, làm giảm đáng kể proline trong cây chuyển nạp gen. Sự tích lũy hàm lượng proline là minh chứng bổ sung trong khi bị stress áp suất thẩm thấu và thể hiện khả năng đáp ứng của cây.

Gen *P5CR* mã hóa proline, hàm lượng proline giảm theo sự tăng nhiệt độ vì promoter nhạy cảm với sốc nhiệt và việc thiết kế “antisense” trong những cây biến nạp (de Ronde và ctv. 2000).



Hình 4: Ảnh hưởng của nhiệt độ và khô hạn trên sự tổng hợp proline, mô tả thao tác của IHSP trên cây T1 transformants. Cả những cây transformants và cây đối chứng được cho vào nghiệm thức stress khô hạn và nhiệt độ 40⁰C (de Ronde và ctv. 2000).

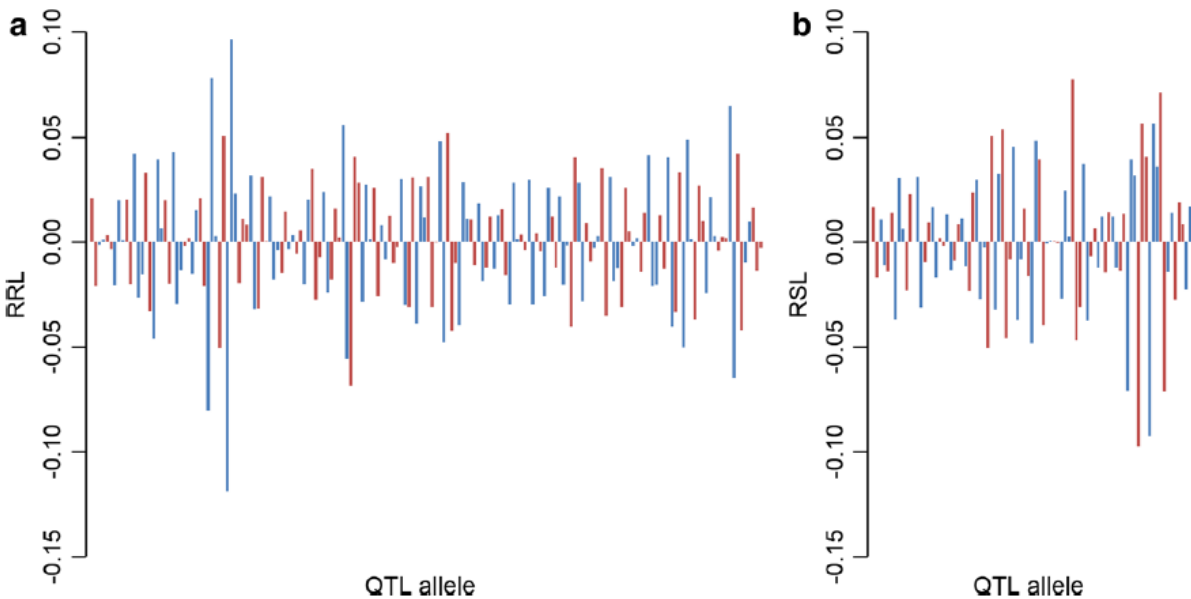
Công trình của Wang và ctv. (2018): Họ gen LEA (late embryonic abundant) có trong phôi các loài thực vật bậc cao, có nhiệm vụ bảo vệ phôi tránh khỏi tổn thương do stress phi sinh học, đặc biệt là khô hạn và mặn. Theo nghiên cứu này, gen *GmLEA2-1* được dòng hóa (cloned) từ lá cây đậu nành thông qua xử lý 10% polyethylene glycol 6000 (PEG6000). Kết quả chạy pRT-PCR (quantitative real-time PCR) cho thấy sự thể hiện phổ gen *GmLEA2-1* ở những mô khác nhau (rễ, thân, lá, bông, quả đậu, phôi non và phôi trưởng thành). Gen *GmLEA2-1* có mức độ đồng dạng về trình tự gen khá thấp với các gen điển hình LEA trong cùng một nhóm từ nhiều loài thực vật khác nhau, nhưng đều có những chức năng rất giống nhau. Sự thể hiện mạnh mẽ gen *GmLEA2-1* trong cây *Arabidopsis thaliana* chuyển nạp gen, cho thấy tính chống chịu với khô hạn và mặn của chúng. Khối lượng sinh khối tươi và khô của cây non, chiều dài rễ chính và mật số rễ ngang của cây biến nạp gen *Arabidopsis* đều cao hơn cây *Arabidopsis* nguyên thủy (WT) dưới điều kiện bị stress khô hạn và mặn. Các nguyên tố *Cis*-acting trong promoter của gen *GmLEA2-1* cũng được xác định. Số liệu này chứng minh rằng protein GmLEA2-1 có vai trò quan trọng trong tính chống chịu mặn và khô hạn của cây trồng (Wang và ctv. 2018).

Xem <https://www.ncbi.nlm.nih.gov/pubmed/30257578>

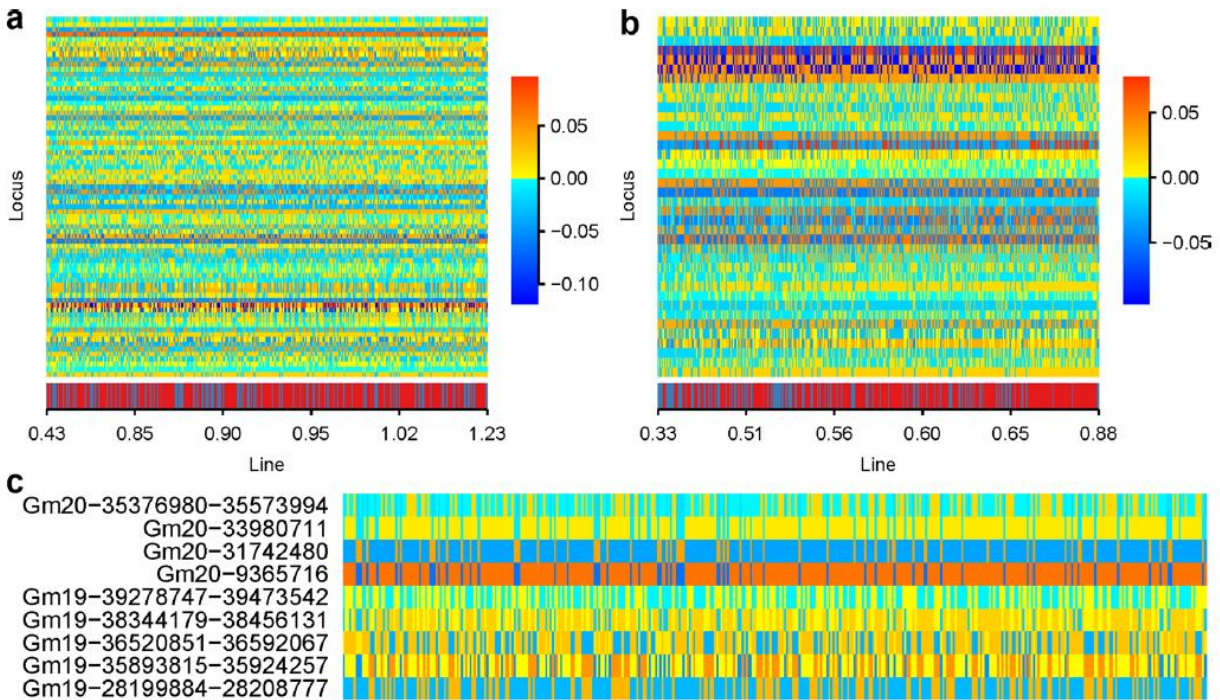
Công trình của Khan và ctv. (2018): RTM-GWAS đã xác định 111 DT QTLs, 262 alleles với tần suất cao QEI và biến dị di truyền đạt 88.55-95.92% PV trong quần thể NAM, từ đó, người ta xây dựng các ma trận QTL-allele và những gen ứng cử viên đã được chú thích. Chống chịu khô hạn (DT) là một những thách thức lớn nhất của nhà chọn giống đậu nành. Quần thể NAM (nested association mapping: bản đồ phối hợp lồng ghép) với tổng số 403 dòng đậu nành phát xuất từ quần thể cận giao tái tổ hợp (RIL) của cặp lai: M8206 × TongShan và ZhengYang × M8206 được trắc nghiệm tính trạng chịu hạn với nghiệm thức xử lý bằng polyethylene-glycol (PEG) trong vụ đậu mùa xuân và mùa hè. Quần thể này được chạy trình tự DNA với kỹ thuật “restriction-site-associated DNA sequencing” (viết tắt là RAD-seq), cho sàng

lọc với phần mềm “minor allele frequency” (MAF) ≥ 0.01 , có tất cả 55.936 chỉ thị SNPs được ghi nhận và được tổ chức thành 6137 SNP LD blocks (linkage disequilibrium blocks: viết tắt là SNPLDBs). Nghiên cứu đa loci với hai giai đoạn có hạn chế theo kiểu “genome-wide association” (viết tắt là RTM-GWAS) đã xác định được 73 QTLs với 174 alen; và 38 QTLs với 88 alen đóng góp cho ảnh hưởng chính thức là 40,43 và 26.11% đối với phương sai kiểu hình (PV) và tương tác giữa QTL với môi trường (QEI). Kết quả ghi nhận 24.64 và 10.35% đối với phương sai kiểu hình cho tính trạng RRL (relative root length) và RSL (relative shoot length), theo thứ tự. Tính trạng chống chịu khô hạn được xác định là có tỷ lệ QEI biến thiên mạnh mẽ (37.52-41.65%), cộng thêm biến thiên di truyền là 46.90-58.40% trong phương sai kiểu hình 88.55-95.92%. Những QTLs-alen được xem xét dưới hai khía cạnh: ảnh hưởng chính và ảnh hưởng QEI theo ma trận “QTL-allele”, kết quả hiển thị được kiến trúc di truyền và QEI của 3 giống bố mẹ và con lai trong quần thể NAM. Từ những ma trận ấy, giống tốt nhất có thể được dự đoán sẽ có giá trị trung bình hơn hẳn hai chỉ thị (indicators) (WAV) ở giá trị 1,873, trong khi cặp lai “top ten” giữa những dòng RILs này đạt thứ hạng 95 “percentile WAV” với giá trị 1,098-1,132, vượt trội so với trung bình của bố mẹ (0,651-0,773), nhưng thấp hơn giá trị 1,873. Điều này cho thấy tiềm năng của thao tác chồng gen (quy tụ gen) với thuật ngữ quốc tế là “pyramiding”. Từ những ma trận nói trên, có 134 gen ứng cử viên đã được chú thích (annotated) thông qua 9 tiến trình sinh học. Đây là phương pháp tiếp cận mới phục vụ chọn giống ở mức độ phân tử trên cơ sở “QTL-allele-based genomic selection” để có những dòng được chọn ưu việt như mong muốn (Khan và ctv. 2018).

Xem <https://www.ncbi.nlm.nih.gov/pubmed/29980855>



Hình 2: Ảnh hưởng của alen tác động QTLs tính trạng chống chịu khô hạn (DT) đối với chiều dài rễ tương đối RRL (a), chiều dài chồi thân tương đối RSL (b) của quần thể đậu nành NAM. Màu đỏ và xanh chỉ ra những QTL khác nhau. Mỗi thanh đặc trưng cho một alen trên một QTL với ảnh hưởng âm tính và dương tính. Có 78 QTL với 174 alen cho tính trạng RRL và 38 QTL với 88 alen cho tính trạng RSL (Khan và ctv. 2018)



Hình 3: “QTL–allele matrices” của tính trạng RRL và RSL trong quần thể đậu nành NAM. **(a)** QTL–allele matrix của RRL. **(b)** QTL–allele matrix của of RSL. **c** Upper part (9 QTLs) của **a** với 9 markers và số tế bào chứa alen của chúng. Trục X của mỗi ma trận biểu thị cho giá trị tính trạng của dòng đậu nành, màu đỏ thuộc về MT và màu xanh thuộc về ZM. Trục Y biểu thị QTLs của tính trạng có ảnh hưởng tác động của alen biểu hiện trong tế bào nhuộm màu, “màu nóng” là những “positive alleles”, “màu lạnh” là những “negative alleles”, độ sâu của màu sắc cho thấy mức độ tác động bởi alen. Tên QTL/marker bị bỏ sót trong mục **a, b** do giới hạn về không gian. Nội dung **c**, phần trên của **a** với các tế bào chứa alen được nói rộng để trình bày tên các marker cho thấy sự phân bố của những alen trong các dòng đậu nành (Khan và ctv. 2018).