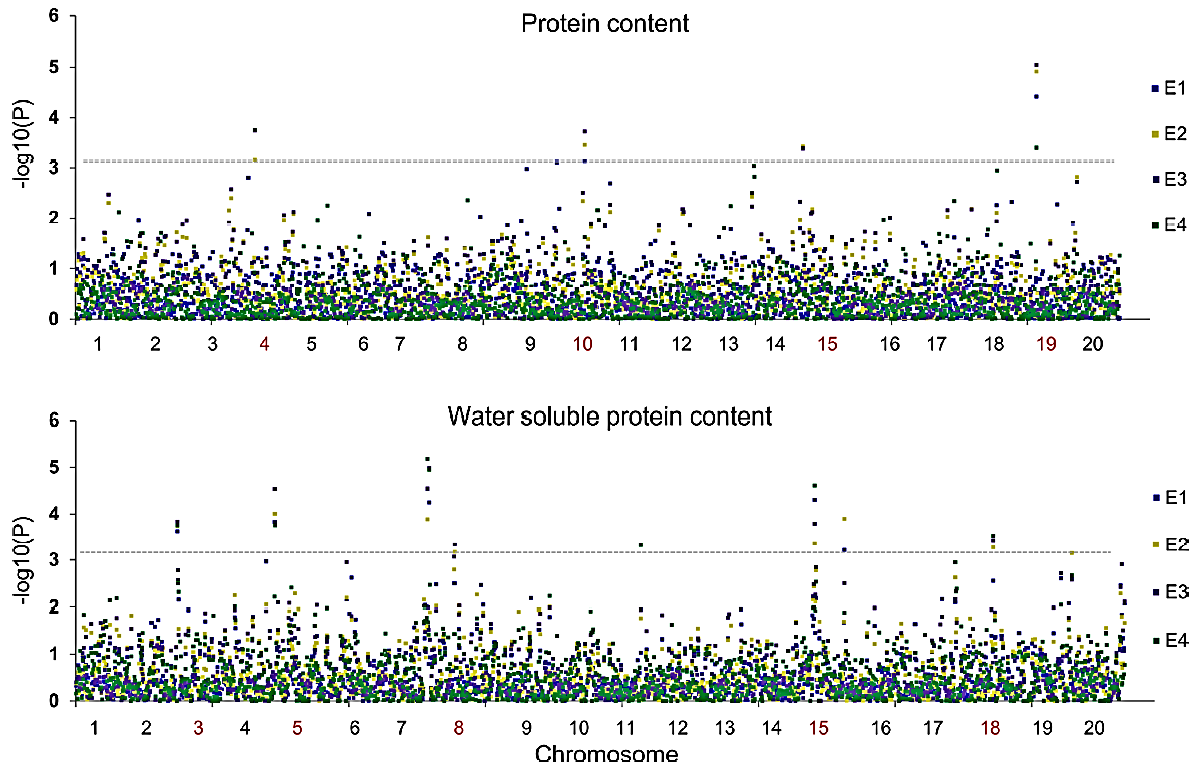


DI TRUYỀN PHẨM CHẤT HẠT

DI TRUYỀN HÀM LƯỢNG PROTEIN TAN TRONG NƯỚC

Zhang và ctv. (2014) đã thực hiện công trình nghiên cứu khoa học về protein hòa tan (water-soluble protein content: viết tắt là SPC). Bốn loci chủ lực SPC đã được xác định, giải thích được 8.5–15.1 % biến thiên kiểu hình. Kết quả giải thích tại sao giống đậu nành bất kỳ nào đó có hàm lượng protein (PC) cao nhưng SPC thấp.

Tính trạng SPC (water-soluble protein content) là yếu tố cực trọng xét trên phương diện dinh dưỡng và sản lượng protein của đậu nành. Tuy nhiên, có ít cơ sở dữ liệu sẵn sàng cho người ta tahm khảo về bản chất di truyền của nó và cơ chế làm cho tăng thêm SPC. Trong nghiên cứu của Zhang và ctv. (2014), người ta sử dụng tập đoàn giống đậu nành gồm 192 mẫu giống từ nhiều vùng địa lý rất xa nhau để xác định trên hệ gen cây đậu nành vùng gen đích liên quan đến tính trạng PC và SPC. Người ta sử dụng bộ chỉ thị 1.536 SNP makers và 232 haplotypes. Panel rất đa dạng này có biến thiên di truyền vô cùng lớn đối với tính trạng PC và SPC. Thực hiện “association mapping” với 3 phương pháp tối thiểu hóa những phối hợp dương tính giả. Kết quả trong 4/8 SNPs và 3/6 haplotypes là có sự phối hợp rất có ý nghĩa của PC/SPC giống đậu nành từ hai hoặc nhiều ngoại cảnh khác biệt nhau trên cơ sở thuật toán “mixed model”. Một SNP phối hợp có ý nghĩa cao với PC, BARC-021267-04016, định vị tại 0.28 cM cách với gen điều khiển glycinin, *G7*, và được phát hiện có ở 4 môi trường khác nhau. Bốn loci đặc biệt SPC, đó là BARC-029149-06088, BARC-018023-02499, BARC-041663-08059 và haplotype 15 (hp15), được xác định khá ổn định trên nhiễm sắc thể 5 và 8. Điều đó giải thích được 8.5–15.1 % biến thiên kiểu hình. Hơn nữa, gen điều khiển glutelin type-B2-like được xác định trên nhiễm sắc thể 8 và có thể có liên quan với sự hòa tan của protein. Những markers như vậy, định vị tại những QTL đã báo cáo trước đây, tái xác định rằng những nghiên cứu trước đó chính xác và xác định các gen có liên quan đến protein. Những chỉ thị mới về SNPs và haplotypes rất quan trọng để hiểu được cơ sở di truyền của tính trạng PC và SPC. Thêm vào đó, khi so sánh tương quan và loci di truyền của hai tính trạng PC và SPC, người ta cung cấp kiến thức mới làm thế nào giống đậu nành có hàm lượng protein cao mà hàm lượng SPC thấp (Zhang và ctv. 2014).



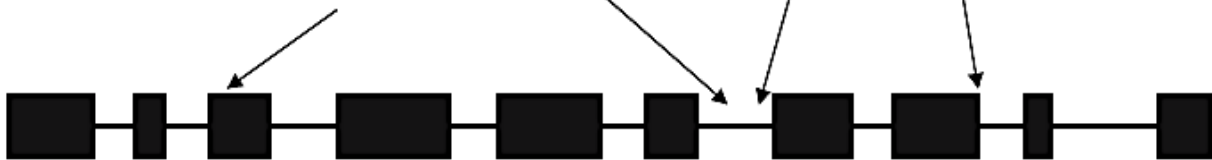
Hình 12: kết quả phân tích GWAS (genome-wide association mapping) tính trạng PC và SPC trên cơ sở SNPs ở 4 môi trường khác nhau (E1–E4). Nhiễm sắc thể được đánh dấu màu đỏ biểu thị SNP clusters phối với tính trạng PC hoặc SPC được xác định ít nhất trên 2 môi trường. Vạch ngang biểu thị đường chuẩn giá trị $-\log P \geq 3.11$ (Zhang và ctv. 2014).

DI TRUYỀN HÀM LƯỢNG PHYTATE THẤP

Yuan và ctv. (2007) đã thực hiện công trình nghiên cứu về di truyền hàm lượng phytate thấp trong cây đậu nành. Phytic acid (PA, *myo*-inositol 1, 2, 3, 4, 5, 6 hexakisphosphate) rất quan trọng đến phẩm chất dinh dưỡng của b khô đậu nành. Organic phosphorus (P) trong PA không thể tiêu hóa được cho người và gia súc không phải loài nhai lại, gây ra ô nhiễm môi trường do P trong nước ngấm và trong phân gia súc thải ra. Hai giống đậu nành đột biến mới [*Glycine max* L. (Merr.)] có *low phytic acid (lpa)* đã được xác định và định tính rõ ràng. *Gm-lpa-TW-1* có sự suy giảm phytic acid P (PA-P) đến 66.6% và làm tăng gấp 6 lần lân vô cơ (Pi), và *Gm-lpa-ZC-2* có sự suy giảm PA-P đến 46.3% và làm tăng gấp 1,4 lần hàm lượng lân vô cơ Pi, so với giống nguyên thủy chưa đột biến gen. Sự suy giảm PA-P và gia tăng Pi trong *Gm-lpa-TW-1* ở mức độ “molar”; sự suy giảm PA-P trong *Gm-lpa-ZC-2*, đi liền với sự tăng Pi và sự thấp hơn hàm lượng inositol phosphates. Trong cả hai dòng đột biến, hàm lượng lân tổng số vẫn giống như giống nguyên thủy. Hai đột biến *lpa* đều có khả năng di truyền được như một gen lặn nhưng không có tính chất “allelic”. Số liệu về trình tự gen cho thấy dòng đột biến *Gm-lpa-TW-1 lpa* có từ 2 bp deletion (mất đoạn) của gen D-*myo*-inositol 3-phosphate synthase (MIPS1 EC 5.5.1.4) - gen 1 (*MIPS1*). Đột biến *lpa* của *Gm-lpa-ZC-2* được ghi trên bản đồ LG B2, liên kết chặt với microsatellite loci Satt416 và Satt168, với khoảng cách di truyền 4.63 -9.25 cM, theo thứ tự. Đột biến này có thể biểu thị một locus *lpa* mới trong hệ gen cây đậu nành. Mật độ hạt của *Gm-lpa-ZC-2* giống như dòng nguyên thủy. Tuy nhiên, *Gm-lpa-TW-1* có *signiWcantly* bị suy

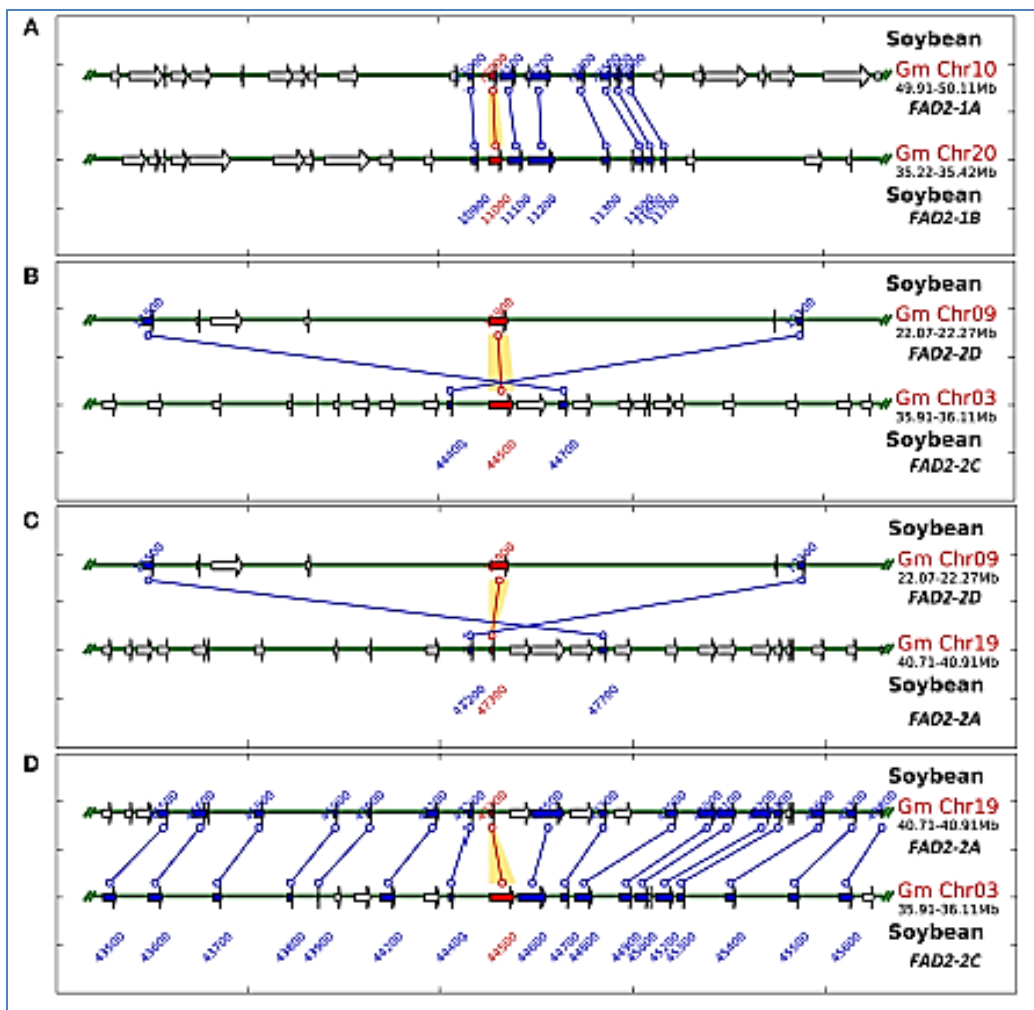
giảm khi trồng đậu nành ở vùng cận nhiệt đới. Khảo nghiệm trên đồng ruộng các dòng đột biến này cho thấy đột biến *lpa* của giống *Gm-lpa-ZC-2* không làm năng suất giảm. Đột biến gen *lpa* của *Gm-lpa-ZC-2*, liên kết với chỉ thị phân tử SSR sẽ phục vụ có hiệu quả cho chương trình cải tiến giống đậu nành nhờ chỉ thị phân tử (Yuan và ctv. 2007).

Taiwan 75: 459 CAagATTCA.... 1502 GT- TTT....1652 GaC1810 CCcG
Gm-lpa-TW-1: 459 CA --ATTCA.... 1502 GT- TTT....1652 GaC1810 CCcG
 Williams 82: 459 CAagATTCA.... 1502 GTaTTT....1652 GgC1810 CCaG



Hình 14: Kiến trúc của gen *MIPS1* và biến dị alen (Yuan và ctv. 2007).

DI TRUYỀN HÀM LƯỢNG OLEIC ACID



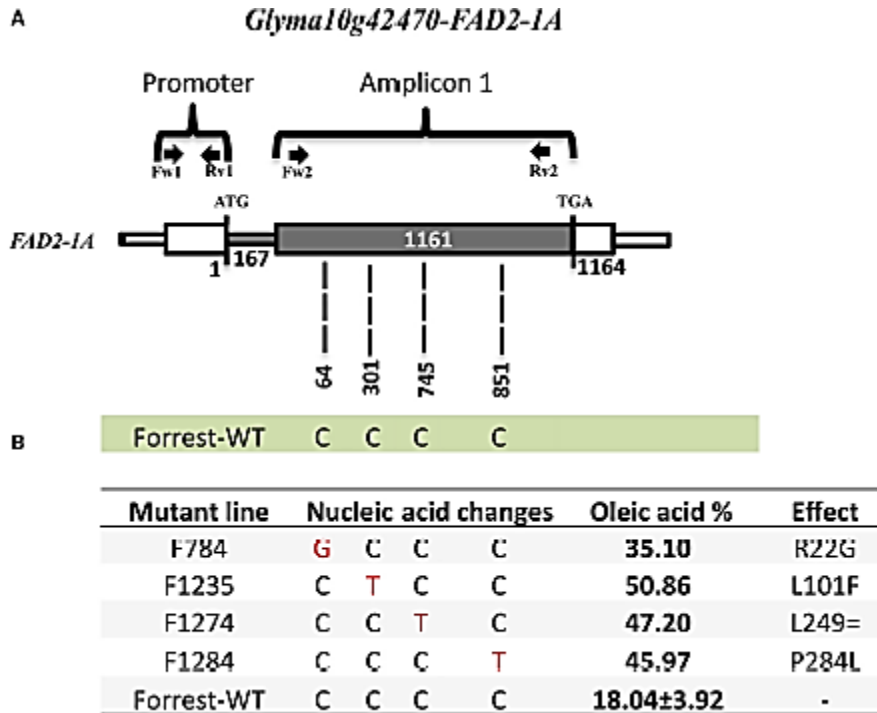
Hình 16: Sơ đồ họ gen **GmFAD2** có những đoạn phân tử lặp đoạn trong hệ gen cây đậu nành (Lakhssassi và ctv. 2017).

Naoufal Lakhssassi và ctv. (2017) đã tiến hành nghiên cứu hàm lượng dầu đậu nành với mức điển hình là 18–20% oleic acid. Việc gia tăng hàm lượng oleic acid rất có lợi cho sức khỏe người tiêu dùng và sản xuất biodiesel. Đột biến các gen *FAD2-1* đã được báo cáo trên tạp chí *Frontier in Plant Science* (13-3-2017) làm tăng hàm lượng oleic acid. Tập đoàn dòng đậu nành đột biến với 1.037 mutants từ giống đậu nành Forrest đã được sàng lọc bằng phương pháp nghiên cứu di truyền ngược (TILLING) nhằm xác định các đột biến xảy ra trong gen *FAD2*. Không có *fad2* mutants được xác định bằng “gel-based TILLING”, bèn gen *fad2-1A* và một gen đột biến *fad2-1B* được xác định có hàm lượng oleic acid trong hạt cao theo kết quả sàng lọc / di truyền thuận (forward genetic) và đọc trình tự “subsequent target”. Phương pháp TILLING được sử dụng thành công như một cách tiếp cận có tính chất “non-transgenic” của di truyền ngược để xác định các đột biến gen điều khiển tính trạng nông học quan trọng này. Tuy nhiên, kỹ thuật ấy vẫn có những hạn chế nhất định ví dụ như thành phần của dầu do số bản sao chép gen và tính chất tương đồng trong hệ gen cây đậu nành. Trong cây đậu nành, gen *FAD2* biểu hiện có hai bản sao chép, *FAD2-1* và *FAD2-2*. Hai thành phần *FAD2-1*: *FAD2-1A* và *FAD2-1B*; ba thành phần *FAD2-2*: *FAD2-2A*, *FAD2-2B*, và *FAD2-2C* đã được các nhà nghiên cứu công bố. Tính chất syntenic, phylogenetic, và phân tích *in silico* cho thấy hai thành viên bổ sung của họ gen *FAD2* là *GmFAD2-2D* định vị trên nhiễm sắc thể 9 và *GmFAD2-2E* trên NST 15. Họ dự đoán các thành viên của *FAD2-2* trên NST 19 là *GmFAD2-2A* và *GmFAD2-2B*, và 03 (*GmFAD2-2C*) (Lakhssassi và ctv. 2017).

Một công trình khác rất đáng chú ý là công trình của Kulkarni và ctv. (2018) tại ĐH Missouri, xem xét mối tương quan giữa protein và hàm lượng dầu như sau:

Nguồn: [Kulkarni KP](#), [Patil G](#), [Vallivodan B](#), [Vuong TD](#), [Shannon JG](#), [Nguyen HT](#), [Lee JD](#). 2018. Comparative genome analysis to identify SNPs associated with high oleic acid and elevated protein content in soybean. [Genome](#). 2018 Mar;61(3):217-222.

Các giống đậu nành có hàm lượng oleic acid cao và hàm lượng protein cao (HOEP) được lai với 5 giống cao sản nhưng hàm lượng oleic acid thấp, protein trung bình (NOAP). Cá thể được chọn lọc mang đi khảo nghiệm tại 6 vùng khác nhau thuộc ba địa điểm trồng đậu nành; đánh giá kiểu hình về tính trạng protein, dầu, các thành phần của acid béo. Hàm lượng protein trung bình của bố mẹ, các dòng HOEP, và NOAP đạt 34.6%, 38%, và 34.9%, theo thứ tự. Hàm lượng oleic acid của bố mẹ, dòng HOEP, và NOAP đạt 21.7%, 80.5%, và 20.8%, theo thứ tự. Dòng cây HOEP mang cả gen đột biến *FAD2-1A* (S117N) và *FAD2-1B* (P137R) làm cho kiểu hình cây thuộc nhóm có hàm lượng oleic acid cao. Phân tích theo kiểu genome so sánh, với số liệu “resequencing” toàn bộ hệ gen cây đậu nành, xác định được 6 gen mang chỉ thị SNP phối hợp có ý nghĩa về thống kê với tính trạng đang phân tích. Một chỉ thị SNP tại vị trí gen giả định **Glyma.10G275800** liên quan đến hàm lượng protein cao, và hàm lượng palmitic, oleic, cũng như linoleic acids. Những gen này từ quãng giữa các marker mà phân tích QTL trước đây đã kết luận lại không mang chỉ thị SNPs có liên quan đến hàm lượng protein và thành phần của acid béo trong những dòng đậu nành đang nghiên cứu. Điều này cho thấy tất cả những gen như vậy ngoại trừ Glyma.10G278000 có thể là những gen mới kết hợp với những tính trạng nêu trên (Kulkaki và ctv. 2018).



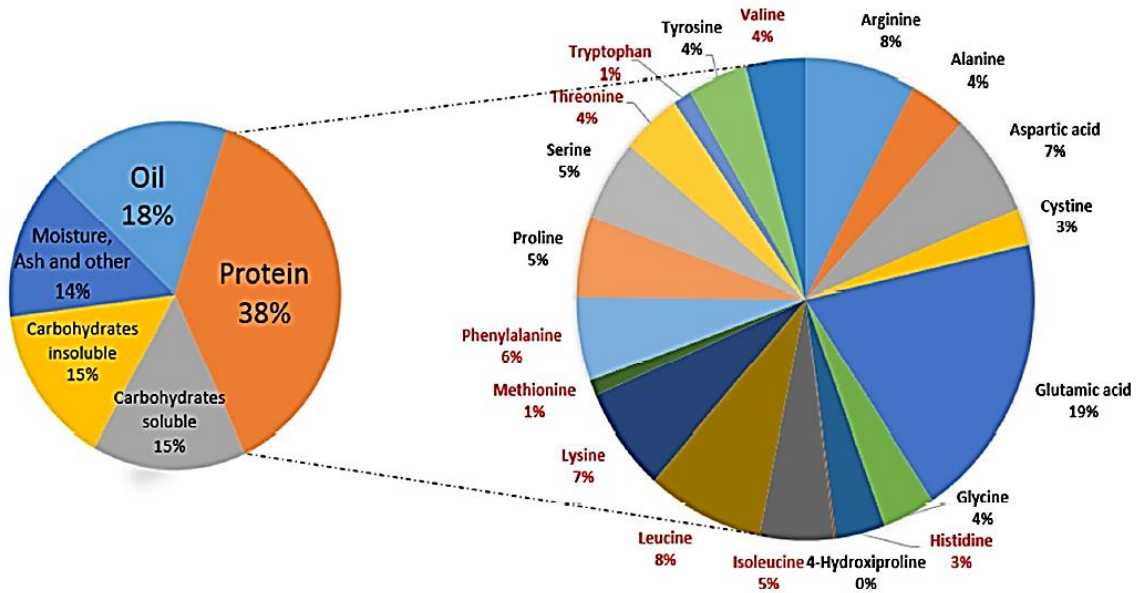
Hình 17: Những đột biến *FAD2-1A* được sàng lọc đối với 4 mutants có hàm lượng oleic acid cao. (A) Mô phỏng gen *FAD2-1A* của giống gốc Forrest. (B) Protein *FAD2-1A* dự đoán cho thấy có khác biệt về amino acid giữa giống Forrest-WT và 4 dòng đột biến có hàm lượng oleic cao theo phân tích di truyền thuận (forward genetics), 3 đột biến kiểu “missense” (R22G, L101F, và P284L), một đột biến “silent” (L249=) trong gen *FAD2-1A*. Hàm lượng oleic acid biểu thị mức độ cao nhất được ghi nhận trong mỗi dòng. Primers cho TILLING và chạy trình tự gDNA được biểu thị bằng mũi tên (Lakhssassi và ctv. 2017).

DI TRUYỀN HÀM LƯỢNG PROTEIN HẠT

Cải tiến di truyền hàm lượng protein hạt là một công trình nghiên cứu rất phức tạp vì có sự tương quan nghịch với hàm lượng dầu, năng suất và nhiệt độ. Patil và ctv. (2017) đã tổng quan các kết quả nghiên cứu này thể hiện qua bản đồ di truyền, hệ gen học, những lỗ hổng trong kiến thức, và sơ lược nhu cầu của những phương pháp nghiên cứu tổng hợp.

Protein của bánh dầu đậu nành có từ hạt đậu nành [*Glycine max* (L) Merr.] là thực phẩm chính cho chăn nuôi gia cầm và gia súc khác. Protein là yếu tố then chốt xác định giá trị dinh dưỡng và giá trị kinh tế của giống đậu nành. Cải tiến di truyền hàm lượng protein trong hạt đậu nành là mục tiêu mà các nhà chọn giống đều mong muốn. Người ta xác định những QTLs (quantitative trait loci) điều khiển hàm lượng protein và hình thành bản đồ QTL trên nhiễm sắc thể 20 (*LG-I*), và 15 (*LG-E*). Tuy nhiên, thực tế trong chọn giống đậu nành luôn phải đối mặt với thách thức, vì protein trong hạt có tương quan nghịch với năng suất hạt, và những thành phần khác có trong hạt như dầu và sucrose. Tương tác giữa giống và môi trường (nhiệt độ trong khi phát triển hạt) cũng là thách thức cho nhà chọn giống. Patil và ctv. (2017) đã đề cập đến những yếu tố hạn chế có liên quan đến hàm lượng protein trong hạt đậu nành và phẩm chất dinh dưỡng, những yếu tố tiềm năng điều hòa hàm lượng protein dự trữ trong hạt. Họ đã mô tả những thuận lợi của phương pháp NGS (next-generation sequencing) để phát hiện chính xác những “variants” tự nhiên và sự tích

hợp của chúng vào kỹ thuật đánh giá kiểu gen theo truyền thống và theo kỹ thuật chính xác. Phân tích “synteny” của QTL trên nhiễm sắc thể 15 và 20 đã được hoàn thiện. Họ đã đề cập đến phương pháp tiếp cận mới về tích hợp QTL điều khiển protein và amino acid, GWAS (genome-wide association studies), kỹ thuật tái đọc trình tự trên toàn hệ gen (hole-genome resequencing), và xây dựng cơ sở dữ liệu “transcriptome” để xác định những điểm nóng trong hệ gen đậu nành phục vụ cho sự dung hợp alen (allele introgression) và cải tiến được protein trong khô dầu đậu nành (Patil và ctv. 2017).



Hình 9: Thành phần amino acid của protein đậu nành. *Màu đỏ* là amino acid cần thiết (EAA), *màu đen* là amino acid không cần thiết. Nguồn: Asif and Acharya (2013), Berk et al. (1992), Kuiken et al. (1949).