

# XÁC ĐỊNH CÁC QTL CHÍNH CHO KÍCH THƯỚC HẠT ĐẬU TƯƠNG VÀ TÍNH TRẠNG TRỌNG LƯỢNG HẠT BẰNG CÁCH SỬ DỤNG QUẦN THỂ RIL TRONG CÁC MÔI TRƯỜNG KHÁC NHAU

Shilin Luo<sup>1,2,3</sup>, Jia Jia<sup>1,2,3</sup>, Riqian Liu<sup>1,2,3</sup>, Ruqian Wei<sup>1,2,3</sup>, Zhibin Guo<sup>1,2,3</sup>, Zhandong Cai<sup>1,2,3</sup>,  
Bo Chen<sup>1,2,3</sup>, Fuwei Liang<sup>1,2,3</sup>, Qiuju Xia<sup>4</sup>, Hai Nian<sup>1,2,3</sup> and Yanbo Cheng<sup>1,2,3\*</sup>

*Võ Như Cẩm biên dịch.*

1. Phòng thí nghiệm trọng điểm cấp nhà nước về bảo tồn và sử dụng tài nguyên sinh học nông nghiệp cận nhiệt đới, Đại học Nông nghiệp Nam Trung Quốc, Quảng Châu, Quảng Đông, Trung Quốc.
2. Phòng thí nghiệm trọng điểm về nhân giống phân tử thực vật của tỉnh Quảng Đông, Đại học Nông nghiệp, Đại học Nông nghiệp Nam Trung Quốc, Quảng Châu, Quảng Đông, Trung Quốc.
3. Phòng thí nghiệm Quảng Đông cho Nông nghiệp hiện đại Lĩnh Nam, Quảng Châu, Quảng Đông, Trung Quốc.
4. Viện chọn giống phân tử lúa gạo, Granlux Associated Grains, Thâm Quyển, Quảng Đông, Trung Quốc.

**Giới thiệu:** Trọng lượng hạt đậu tương [*Glycine max* (L.) Merr.] là một trong những tính trạng chính quyết định năng suất đậu tương và có quan hệ mật thiết với kích thước hạt. Tuy nhiên, cơ sở di truyền của quy định tổng hợp các tính trạng liên quan đến năng suất đậu tương là không rõ ràng.

**Phương pháp:** Để hiểu cơ sở di truyền phân tử cho sự hình thành các tính trạng năng suất đậu tương, nghiên cứu hiện tại tập trung vào lập bản đồ QTL cho các tính trạng kích thước và trọng lượng hạt trong các môi trường khác nhau và khai thác gen mục tiêu.

**Kết quả:** Tổng cộng có 85 QTL liên quan đến các tính trạng về kích thước và trọng lượng hạt được xác định bằng cách sử dụng quần thể dòng lai tái tổ hợp (RIL) được phát triển từ Guizao1×B13 (GB13). Chúng tôi cũng phát hiện 18 QTL ổn định với môi trường. Trong số này, *qSL-3-1* là một QTL mới với hiệu ứng chính ổn định liên quan đến chiều dài hạt. Nó được phát hiện trong tất cả các môi trường, ba trong số đó giải thích hơn 10% phương sai kiểu hình (PV), với mức tối đa là 15,91%. Ngoài ra, *qSW-20-3* là một QTL mới với hiệu ứng chính ổn định liên quan đến chiều rộng của hạt, được xác định trong bốn môi trường. Và lượng phương sai giải thích kiểu hình (PVE) thay đổi từ 9,22 đến 21,93%. Năm nhóm QTL liên quan đến cả kích thước hạt và trọng lượng hạt được tóm tắt bằng nhận dạng nhóm QTL. Mười lăm gen ứng cử viên có thể liên quan đến việc điều chỉnh kích thước và trọng lượng hạt đậu tương cũng được sàng lọc dựa trên chú thích chức năng gen và phân tích làm giàu GO.

**Thảo luận:** Các kết quả cung cấp một tài liệu tham khảo cơ bản về mặt sinh học để hiểu sự hình thành các tính trạng về kích thước và trọng lượng của hạt đậu tương.

## 1. GIỚI THIỆU

Đậu nành là một trong những loại cây lương thực được đánh giá cao trên toàn cầu với nhiều chất dinh dưỡng bao gồm protein, carbohydrate, lipid, khoáng chất, vitamin và các hoạt chất sinh học như isoflavone, saponin, sterol và phospholipid. Với những giá trị quan trọng trong nông nghiệp và kinh tế, đậu tương đóng vai trò quan trọng trong các hoạt động nông nghiệp và công nghiệp như chế biến thực phẩm và thức ăn chăn nuôi, sản xuất dệt may. Trong khi nhu cầu về đậu tương đang gia tăng trên toàn cầu, việc nâng cao

năng suất đậu tương hiện đang nhận được sự quan tâm đáng kể vì tiềm năng nâng cao năng suất của nó.

Kích thước hạt và trọng lượng hạt là những tính trạng số lượng được kiểm soát bởi nhiều gen và bị hạn chế bởi các yếu tố môi trường. Sự hình thành của cả hai tính trạng là một quá trình sinh học phức tạp (Li và cs, 2020). Kích thước hạt thường đo chiều dài hạt, chiều rộng hạt và độ dày hạt (Salas và cs, 2006). Nhiều nghiên cứu đã được thực hiện trên toàn thế giới về nội địa hóa QTL về kích thước và trọng lượng hạt ở đậu tương. Ngược lại, các nghiên cứu về nội địa hóa QTL đối với kích thước hạt tương đối ít hơn so với trọng lượng hạt. Ví dụ: trên trang web dữ liệu về đậu tương SoyBase (<http://www.soybase.org/>) kể từ tháng 3 năm 2022, số ấn phẩm QTL là 18 về chiều dài hạt, 16 về chiều rộng hạt và 23 về độ dày của hạt so với số của các ấn phẩm QTL về trọng lượng hạt là 300.

Salas và cs (2006) đã phát hiện tổng cộng 19 QTL liên kết với hình dạng hạt bằng cách sử dụng ba quần thể RIL trong hai môi trường. Xu và cs (2011) đã sử dụng MAJ Multi-QTL Joint Analysis (MAJ) và lập bản đồ khoảng tổng hợp (CIM) để lập bản đồ các QTL. Tổng cộng có 121 QTL chính được phát hiện trong các quần thể  $F_{2:3}$ ,  $F_{2:4}$  và  $F_{2:5}$  từ các phép lai trực tiếp và tương hỗ của Lishuizhongzihuang với Nannong493-1. Sun và cs (2021) đã sử dụng quần thể RIL để phát hiện các QTL cho các tính trạng kích thước hạt trong bốn môi trường. 10 QTL kiểm soát các tính trạng liên quan đã được xác định, trong đó 5 QTL phân bố trên các nhiễm sắc thể 02, 04, 06, 13 và 16 được phát hiện trong ít nhất 2 môi trường có PVE dao động từ 3,6 – 9,4%. Kết quả của các nghiên cứu tương tự (Salas và cs, 2006; Xu và cs, 2011; Hu và cs, 2013; Niu và cs, 2013; Yang và cs, 2017; Teng và cs, 2018; Gao và cs, 2019; Cui và cs, 2020; Sun và cs, 2021; Kumawat và Xu, 2021) khác nhau ở các phương pháp nội địa hóa khác nhau và nền tảng di truyền khác nhau.

Nghiên cứu đầu tiên về nội địa hóa QTL trọng lượng hạt trong hạt đậu được báo cáo bởi Sax (1923), trong đó ông sử dụng màu sắc của hạt làm marker. Với sự tiến bộ của công nghệ chỉ thị phân tử, các học giả đã phát triển nhiều chỉ thị phân tử khác nhau như đa hình chiều dài đoạn giới hạn (RFLP), lặp lại trình tự đơn giản (SSR) và đa hình đơn nucleotide (SNP) với các quần thể lập bản đồ khác nhau để xác định các QTL của trọng lượng hạt đậu tương. Năm 1996, Maughan và cs (1996) đã phát hiện 6 QTL liên quan đến trọng lượng hạt trong 150 quần thể  $F_2$  và 150 dòng  $F_{2:3}$  của chúng bằng cách sử dụng 91 chỉ thị di truyền đa hình bao gồm RFP, RAPD và SSR. Trương và cs (2016) đã thực hiện một nghiên cứu kết hợp trên toàn bộ bộ gen (GWAS) bằng cách sử dụng 309 nguồn tế bào mầm đậu tương với 31045 SNP markers và các QTL liên quan liên quan đến trọng lượng hạt đến nhiễm sắc thể 04 và 19. Các kết quả khác từ các nghiên cứu trước đó (Mian và cs, 1996; Csanadi và cs, 2001; Lee và cs, 2001; Chapman và cs, 2003; Fasoula và cs, 2004; Hyten và cs, 2004; D. Li và cs, 2008; Han và cs, 2012; Pathan và cs, 2013; Kato và cs, 2014; Yan và cs, 2014; Kulkarni và cs, 2017; Liu D. và cs, 2018) không giống nhau, nhưng các QTL này đã đặt nền móng cho chỉ thị phân tử - chọn giống.

Tóm lại, hầu hết các nhà nghiên cứu sử dụng số lượng quần thể lập bản đồ khác nhau để phát hiện QTL trong các môi trường khác nhau. Tuy nhiên, các nghiên cứu sử dụng nhiều quần thể lập bản đồ hơn có ít bối cảnh thử nghiệm hơn, ngược lại với các nghiên cứu sử dụng ít quần thể lập bản đồ hơn có xu hướng có nhiều môi trường thử nghiệm hơn. Trong nghiên cứu này, một quần thể lập bản đồ được sử dụng để phát hiện QTL trong sáu môi trường cho thấy tiêu chuẩn cao hơn để xác định QTL ổn định. Ngoài ra, việc nội địa hóa các QTL về kích thước và trọng lượng hạt đậu tương theo quần thể RIL ở các môi trường

khác nhau và khai thác các gen tác động chính giúp chúng ta hiểu được cơ sở sinh học của quá trình hình thành tính trạng năng suất đậu tương, có ý nghĩa lý luận và thực tiễn rất lớn trong việc định hướng nhân giống thiết kế phân tử và liên tục nâng cao năng suất đậu tương.

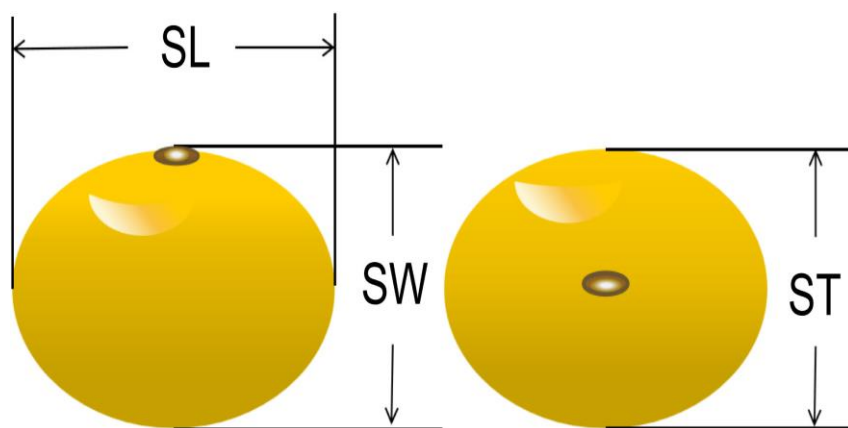
## 2. VẬT LIỆU VÀ PHƯƠNG PHÁP

### 2.1. Vật liệu và cách trồng

Quần thể GB13 RIL, bao gồm 248 dòng, được xây dựng với Guizao1 là mẹ và B13 là bố, được gieo ba lần vào tháng 7/2018, tháng 7/2019 và tháng 7/2020 tại Cơ sở Giảng dạy và Nghiên cứu của Zengcheng (23°14'N, 113°38'Đ), và vào tháng 7/2019 và tháng 7/2020 tại Cơ sở Giảng dạy Thực nghiệm Quảng Châu (23°10'B, 113°22'Đ) (được chỉ định lần lượt là 18ZC, 19ZC, 20ZC, 19GZ và 20GZ). Ngoài ra, một RIL F<sub>1</sub> và 248 F<sub>8:11</sub> của quần thể GB13 RIL đã được phát triển bằng phương pháp gieo một hạt (Jiang và cs, 2019). Các vật liệu thí nghiệm được trồng theo kiểu hoàn toàn ngẫu nhiên với mỗi dòng một hàng, chiều dài hàng 2,0m, khoảng cách hàng 0,5m và khoảng cách cây 0,1m. Việc quản lý đồng ruộng được thực hiện theo các tiêu chuẩn thông thường và không có dịch bệnh hoặc sâu bệnh xảy ra trong quá trình sinh trưởng.

### 2.2. Thống kê và phân tích kiểu hình

Chiều dài hạt (SL), chiều rộng hạt (SW) và độ dày hạt (ST) của hạt đậu tương được đo bằng thước cặp điện tử (Hình 1). Hai mươi hạt từ mỗi dòng được chọn ngẫu nhiên để đo kích thước hạt, sau đó chúng được cân trên cân điện tử. Sau ba lần lặp lại, năm lần trọng lượng trung bình được sử dụng làm dữ liệu kiểu hình cho tính trạng trọng lượng hạt (trọng lượng 100 hạt, HSW).



**Hình 1.**  
Đo chiều dài hạt (SL),  
chiều rộng hạt (SW)  
và độ dày hạt (ST)

Giá trị trung bình của dữ liệu kiểu hình từ quần thể GB13 RIL trong năm môi trường tự nhiên được tính là dữ liệu kiểu hình của môi trường thứ sáu (môi trường kết hợp, được chỉ định là CE). Thống kê mô tả, phân tích hồi quy tuyến tính và tính toán độ lệch và độ nhọn được thực hiện trên dữ liệu kiểu hình từ sáu môi trường bằng phần mềm GraphPad Prism (GraphPad Software Inc. v.7.0.4, Hoa Kỳ) và các bảng hoặc số liệu tương ứng đã được tạo. Phân tích phương sai (ANOVA) và ước tính hệ số di truyền theo nghĩa rộng ( $h^2$ ) được thực hiện trên dữ liệu kiểu hình từ năm môi trường tự nhiên bằng phần mềm Genstat phiên bản thứ 12. Các công thức (Nyquist và Baker, 1991; Elattar và cs, 2021) như sau:

$$h^2 = \sigma_g^2 / (\sigma_g^2 + \sigma_{ge}^2/n + \sigma_e^2/nr)$$

trong đó  $\sigma_g^2$  biểu thị phương sai kiểu gen,  $\sigma_{ge}^2$  biểu thị phương sai tương tác kiểu gen  $\times$  môi trường và  $\sigma_e^2$  biểu thị sai số phương sai. “n” biểu thị số lượng môi trường và “r” biểu thị số lần lặp lại trên mỗi môi trường.

### 2.3. Nội địa hóa QTL cho kích thước hạt và tính trạng trọng lượng và phân tích

Phương pháp lập bản đồ khoảng tổng hợp (Zeng, 1994) đã được sử dụng để lập bản đồ các QTL với bản đồ liên kết thừng mật độ cao được xây dựng trước đó trong phòng thí nghiệm, như đã mô tả, bản đồ chứa 56.561 SNP và 3715 marker thừng có tổng chiều dài bộ gen là 3049,2cM, với độ dài của nhóm liên kết dao động từ 120,22–211,37cM, Khoảng cách trung bình giữa các marker là 0,80cM và số lượng marker bin trên mỗi nhóm thay đổi từ 147 đến 259. (Jiang và cs, 2019). Hiệu ứng cộng (ADD) và PVE của mỗi QTL được phân tích bằng phần mềm WinQTLCart 2.5. Ngưỡng LOD cho các tính trạng khác nhau được đặt thành 2,5 ở mức ý nghĩa 5% và ngưỡng LOD lớn hơn 2,5 được coi là cơ sở để xác định sự hiện diện của các QTL. Các QTL nằm trên cùng một nhiễm sắc thể với các khoảng đánh dấu liền kề được kết hợp thành một QTL nếu khoảng tin cậy của chúng trùng nhau. Các QTL được đặt tên theo “q + tính trạng + nhiễm sắc thể + số thứ tự”, với ký hiệu “-” ở giữa (McCouch và cs, 1997).

### 2.4. Xác định các nhóm QTL và QTL ổn định

Nói chung, các QTL được phát hiện trong ít nhất hai môi trường được coi là các QTL ổn định (Qi và cs, 2017), nhưng có sáu môi trường thử nghiệm trong nghiên cứu này, vì vậy các QTL được phát hiện trong ít nhất ba môi trường được coi là các QTL ổn định trong nghiên cứu này để xác định các QTL ổn định.

Tất cả các QTL được sắp xếp theo nhiễm sắc thể là điều kiện chính và vị trí vật lý là điều kiện thứ cấp. Các QTL có nhiễm sắc thể giống hệt nhau ở vị trí vật lý chông chéo hoặc liền kề (dưới 5cM) được nhóm thành một nhóm và được xác định là nhóm QTL nếu nó liên quan đến ít nhất hai tính trạng (Liu và cs, 2017; Wang và cs, 2022). Để khai thác tác dụng chính và các gen ứng cử viên ổn định, các nhóm QTL được xác định chứa ít nhất một QTL ổn định trong nghiên cứu này.

### 2.5. Dự đoán gen ứng cử viên

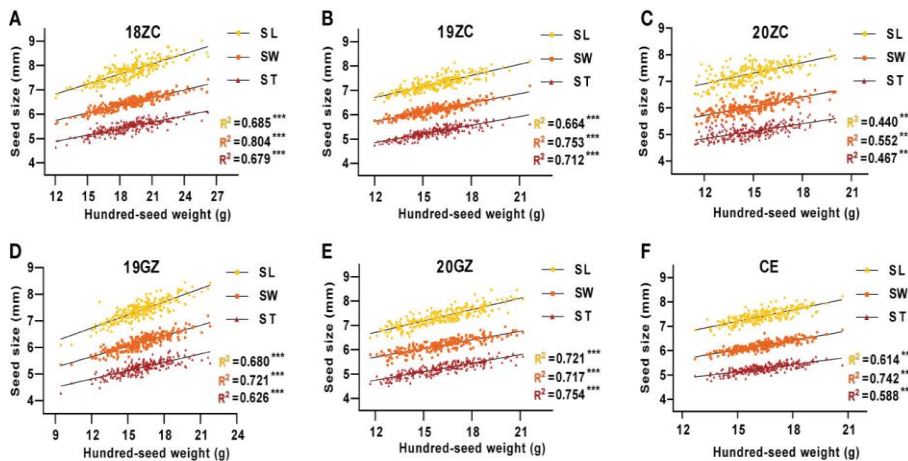
Đầu tiên, phân tích làm giàu thuật ngữ GO (Gene Ontology) đã được thực hiện cho các gen trong khoảng thời gian của các nhóm QTL đã xác định. Điều này rất hữu ích để chọn các gen ứng cử viên dựa trên sự biến đổi trình tự của các gen giữa bố và mẹ. Các chú thích GO đã được tải xuống từ trang web dữ liệu đậu tương SoyBase ([//www.soybase.org/](http://www.soybase.org/)), và sau đó là trang web lập bản đồ phân tích làm giàu gen trực tuyến WEGO2.0 ([//wego.genomics.org.cn/](http://wego.genomics.org.cn/)) được sử dụng để tạo thành phân tử bào làm giàu gen, chức năng phân tử và sơ đồ trực quan hóa các quá trình sinh học. Tiếp theo, các gen được biểu hiện cao hoặc cụ thể trong hạt đậu tương sẽ được sàng lọc thêm. Dữ liệu biểu hiện gen được lấy từ trang web dữ liệu đậu nành SoyBase. Cuối cùng, các gen được chọn đã được chú thích chức năng bởi cơ sở dữ liệu Phytozome ([//phytozome-next.jgi.doe.gov/](http://phytozome-next.jgi.doe.gov/)), và sau đó các gen ứng cử viên được xác định dựa trên tài liệu liên quan và các chức năng gen liên quan. Trong khi đó, bản đồ nhiệt của biểu hiện gen ứng cử viên đã được sử dụng để phân tích biểu hiện cụ thể của mô (Jia và cs, 2022) bởi trang web trực tuyến ([//www.omicshare.com/tools/Home/Soft/heatmap](http://www.omicshare.com/tools/Home/Soft/heatmap)) và cấu trúc của các gen ứng cử viên đã được lập bản đồ bằng trang web trực tuyến của Máy chủ hiển thị cấu trúc gen ([//gsds.gao-lab.org/index.php](http://gsds.gao-lab.org/index.php)). Dữ liệu sắp xếp lại trình tự của các dòng bố mẹ được dùng để so sánh sự biến đổi của các gen ứng cử viên giữa các dòng bố mẹ. Toàn bộ

bộ gen của các dòng bố mẹ được giải trình tự bằng cách sử dụng Illumina HiSeq X Ten p, với độ sâu trình tự trung bình là 8× (Xian và cs, 2022). Dữ liệu giải trình tự chất lượng cao được đánh giá để dự đoán các biến thể cấu trúc gen.

### 3. KẾT QUẢ

#### 3.1. Phân tích tương quan giữa kích thước hạt và trọng lượng hạt

Phân tích hồi quy tuyến tính kích thước hạt với trọng lượng hạt cho thấy SL, SW và ST đều có tương quan thuận và có ý nghĩa cao với HSW ( $p < 0,01$ ), với R-Squared ( $R^2$ ) của phân tích tuyến tính hồi quy nằm trong khoảng từ 0,440–0,804 (Hình 2). Trong 18ZC, 19ZC, 19GZ, 20ZC và CE,  $R^2$  giữa SW và HSW cao nhất lần lượt là 0,804, 0,721, 0,753, 0,552 và 0,742, trong khi ở 20GZ,  $R^2$  giữa ST và HSW cao nhất là 0,754. Do đó, kích thước hạt thực sự có liên quan chặt chẽ đến trọng lượng hạt. 18ZC, 19ZC và CE có  $R^2$  lớn nhất tiếp theo giữa SL và HSW lần lượt là 0,685, 0,680 và 0,614 và  $R^2$  nhỏ nhất giữa ST và HSW lần lượt là 0,679, 0,626 và 0,588. 19GZ và 20ZC có  $R^2$  lớn nhất tiếp theo giữa ST và HSW là 0,713 và 0,467 và  $R^2$  nhỏ nhất giữa SL và HSW là 0,664 và 0,440. 20GZ có  $R^2$  lớn nhất tiếp theo giữa SL và HSW là 0,721 và  $R^2$  nhỏ nhất giữa SW và HSW là 0,717. Ngoài ra, kết quả trên cho thấy mối tương quan giữa kích thước hạt và khối lượng hạt dường như được áp dụng trong chọn tạo giống đồng ruộng làm tham chiếu cho chọn giống năng suất cao.



**Hình 2.** Mối tương quan giữa kích thước hạt và trọng lượng hạt của quần thể RIL GB13 tại Cơ sở Giảng dạy và Nghiên cứu của Zengcheng (ZC) trong (A) 2018, (B) 2019 và (C) 2020, Cơ sở Giảng dạy Thử nghiệm của Quảng Châu (GZ) trong (D) 2019 và (E) 2020 và (F) môi trường kết hợp của chúng (CE). \*\*\*  $p < 0,001$  (thử nghiệm F).  $R^2$  chỉ ra hệ số tương quan.

#### 3.2. Phân tích thống kê mô tả các tính trạng kích thước, trọng lượng hạt

Dữ liệu kiểu hình được phân tích bằng thống kê mô tả và kết quả cho thấy SL, SW, ST (Hình bổ sung 1) và HSW của bố mẹ Guizaol cao hơn so với bố mẹ B13 của giống bố với sự khác biệt đáng kể (Bảng 1), trong đó đặt nền móng cho việc lập bản đồ QTL. Quần thể GB13 RIL có các tính trạng về kích thước và trọng lượng khác nhau đáng kể, cả hai đều cho thấy sự phân biệt quá mức. Hệ số biến thiên (CV) đối với SL dao động từ 4,33–5,25%, đối với SW từ 3,96–4,78%, đối với ST từ 4,71–5,18% và đối với HSW từ 10,68–12,98%. CV của HSW lớn hơn của SL, SW và ST, nhưng chúng tương đối ổn định, điều này cho thấy rằng cả bố và mẹ đều chứa các gen hoạt động trong sự biến đổi kiểu hình.

**Bảng 1.** Phân tích thống kê các tính trạng kích thước và trọng lượng hạt của bố mẹ và các dòng ở các môi trường khác nhau

Tính trạng #	Môi trường a	Bố mẹ <sup>b</sup>		RILs <sup>c</sup>					
		Guizaol <sup>c</sup>	B13 <sup>d</sup>	Khoảng <sup>f</sup>	Trung bình <sup>g</sup>	SD <sup>h</sup>	CV(%) <sup>i</sup>	Skew <sup>j</sup>	Kurt <sup>k</sup>
SL	18ZC	834 ± 0,02	7,40 ± 0,01	6,66 - 9,01	7,77	0,41	5,25	0,12	-0,20
	19GZ	7,40 ± 0,02	6,90 ± 0,02	6,08 - 8,44	7,40	0,38	5,15	-0,19	0,66

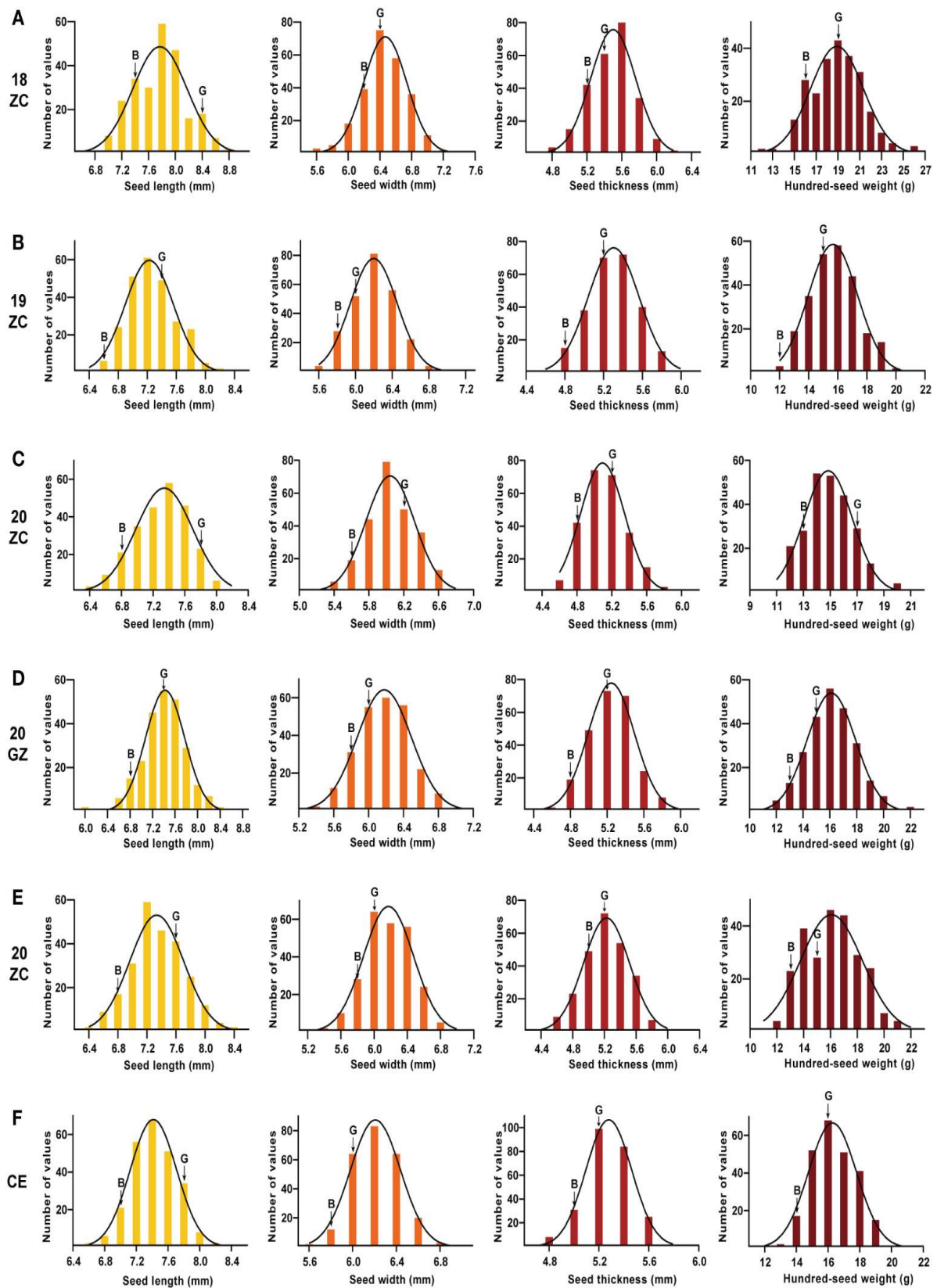
	19ZC	7,40 ± 0,02	6,68 ± 0,02	6,52 - 8,21	7,27	0,32	4,33	0,32	-0,20
	20GZ	7,61 ± 0,01	6,96 ± 0,03	6,48 - 8,46	7,35	0,37	5,04	0,20	0,03
	20ZC	7,87 ± 0,03	6,85 ± 0,01	6,29 - 8,17	7,30	0,35	4,82	-0,35	-0,20
	CE	7,72 ± 0,07	6,94 ± 0,05	6,64 - 8,23	7,42	0,28	3,80	0,07	-0,03
SW	18ZC	6,65 ± 0,01	6,24 ± 0,01	5,59 - 7,43	6,46	0,29	4,44	-0,17	0,40
	19GZ	5,99 ± 0,02	5,84 ± 0,05	5,35 - 7,18	6,18	0,30	4,78	-0,01	0,22
	19ZC	6,02 ± 0,02	5,77 ± 0,02	5,65 - 7,18	6,20	0,25	3,96	0,31	0,46
	20GZ	5,97 ± 0,01	6,08 ± 0,02	5,37 - 6,95	6,18	0,28	4,55	-0,12	-0,23
	20ZC	6,10 ± 0,02	5,68 ± 0,03	5,37 - 6,75	6,05	0,27	4,50	-0,01	-0,27
	CE	6,34 ± 0,11	6,19 ± 0,08	5,66 - 6,85	6,21	0,21	3,39	0,14	-0,21
ST	18ZC	5,53 ± 0,01	5,22 ± 0,01	4,65 - 6,11	5,49	0,26	4,79	-0,22	0,07
	19GZ	5,17 ± 0,02	4,84 ± 0,04	4,28 - 5,90	5,24	0,25	4,83	-0,23	0,60
	19ZC	5,12 ± 0,01	4,90 ± 0,01	4,75 - 5,90	5,30	0,24	4,49	0,01	-0,47
	20GZ	5,11 ± 0,01	5,06 ± 0,01	4,52 - 5,88	5,23	0,27	5,18	-0,14	-0,39
	20ZC	5,22 ± 0,01	4,88 ± 0,02	4,61 - 5,81	5,12	0,24	4,71	0,25	-0,20
	CE	5,16 ± 0,01	5,05 ± 0,05	4,76 - 5,75	5,27	0,18	3,39	-0,10	-0,06
HSW	18ZC	18,85 ± 0,1	16,32 ± 0,06	12,05 - 26,22	18,88	2,45	12,98	0,14	0,28
	19GZ	15,35 ± 0,19	12,90 ± 0,24	9,45 - 21,80	16,09	1,91	11,89	0,02	0,61
	19ZC	14,97 ± 0,06	12,32 ± 0,05	12,00 - 21,63	15,77	1,68	10,68	0,29	0,02
	20GZ	15,40 ± 0,05	13,27 ± 0,13	11,70 - 21,17	16,08	2,01	12,52	0,14	-0,65
	20ZC	17,30 ± 0,12	13,02 ± 0,09	11,43 - 20,03	14,96	1,75	11,71	0,31	-0,17
	CE	16,37 ± 0,30	14,14 ± 0,51	12,70 - 20,70	16,35	1,41	8,62	0,15	-0,34

Ghi chú: # chiều dài hạt (SL), chiều rộng hạt (SW), độ dày hạt (ST) và trọng lượng 100 hạt (HSW); <sup>a</sup> Môi trường; <sup>b</sup> Bố mẹ của quần thể RIL GB13; <sup>c</sup> Giống bố của quần thể RIL GB13; <sup>d</sup> Giống mẹ của quần thể RIL GB13; <sup>e</sup> Dòng lai tái tổ hợp; <sup>f</sup> Khoảng của kích thước hạt và trọng lượng 100 hạt của quần thể RIL GB13; <sup>g</sup> Trung bình của kích thước hạt và trọng lượng 100 hạt của quần thể RIL GB13; <sup>h</sup> Độ lệch chuẩn; <sup>i</sup> CV; <sup>j</sup> Độ lệch; <sup>k</sup> Độ nhọn.

Các giá trị tuyệt đối của độ lệch và độ nhọn của SL, SW, ST và HSW của quần thể GB13 RIL nhỏ hơn 1 (Bảng 1). Ngoài ra, biểu đồ phân bố tần suất (Hình 3) cho thấy dữ liệu kiểu hình về kích thước hạt và các tính trạng trọng lượng thể hiện sự thay đổi liên tục. Các kết quả trên cho thấy kích thước hạt và các tính trạng trọng lượng của quần thể GB13 RIL tuân theo phân bố chuẩn, phù hợp với tính trạng của quần thể RIL và thuộc về tính trạng di truyền số lượng.

### 3.3. Phân tích phương sai và ước tính hệ số di truyền theo nghĩa rộng

Kết quả ANOVA đối với SL, SW, ST và HSW của quần thể GB13 RIL trong năm môi trường tự nhiên (Bảng 2) cho thấy kiểu gen, môi trường và sự tương tác giữa kiểu gen và môi trường có ảnh hưởng đáng kể đến kích thước hạt và các tính trạng trọng lượng của quần thể RIL GB13 ( $p < 0,0001$ ).  $H^2$  đối với các tính trạng kích thước hạt và trọng lượng của quần thể RIL GB13 tương đối cao, dao động từ 0,74–0,83 (Bảng 2). Các tính trạng phù hợp để phân tích nội địa hóa QTL hơn nữa.



**Hình 3.** Phân bố tần số và đường cong phù hợp của các tính trạng về kích thước và trọng lượng hạt đối với quần thể GB13 RIL tại Cơ sở Giảng dạy và Nghiên cứu của Zengcheng (ZC) vào (A) 2018, (B) 2019 và (C) 2020, Cơ sở Giảng dạy Thử nghiệm của Quảng Châu (GZ) trong (D) 2019 và (E) 2020 và (F) môi trường kết hợp của chúng (CE). Chữ “G” biểu thị Guizao1 là mẹ và chữ “B” biểu thị B13 là bố. Tọa độ ngang của các hình có màu khác nhau từ trái sang phải là chiều dài hạt, chiều rộng hạt, độ dày hạt và trọng lượng 100 hạt.

**Bảng 2.** Phân tích phương sai và khả năng di truyền theo nghĩa rộng đối với quần thể RIL GB13 tại năm môi trường tự nhiên.

Tính trạng <sup>#</sup>	Nguồn <sup>a</sup>	Df <sup>b</sup>	SS <sup>c</sup>	MS <sup>d</sup>	P <sub>value</sub> <sup>e</sup>	VC <sup>f</sup>	PV(%) <sup>g</sup>	h <sup>2</sup> (%) <sup>h</sup>
SL	Lặp lại	2	0,74	0,37	<0,0001			0,83
	Kiểu gen	247	294,11	1,19	<0,0001	0,07	46,49	
	Môi trường	4	123,42	30,86	<0,0001	0,04	19,51	
	Tương tác	988	200,95	0,20	<0,0001	0,07	31,77	
	Lỗi	2478	13,38	0,01	<0,0001			
	Tổng biến thể	3719	632,59					
SW	Lặp lại	2	1,09	0,54	<0,0001			0,82
	Kiểu gen	247	164,87	0,67	<0,0001	0,04	45,42	
	Môi trường	4	66,05	16,51	<0,0001	0,02	18,19	
	Tương tác	988	119,22	0,12	<0,0001	0,04	32,84	
	Lỗi	2478	11,80	0,01	<0,0001			
	Tổng biến thể	3719	363,04					
ST	Lặp lại	2	1,09	0,54	<0,0001			0,74
	Kiểu gen	247	118,51	0,48	<0,0001	0,02	38,23	
	Môi trường	4	54,34	13,58	<0,0001	0,02	17,53	
	Tương tác	988	124,30	0,13	<0,0001	0,04	40,1	
	Lỗi	2478	11,72	0,01	<0,0001			
	Tổng biến thể	3719	309,96					
HSW	Lặp lại	2	45,00	22,50	<0,0001			0,75
	Kiểu gen	247	7210,17	29,19	<0,0001	1,46	33,32	
	Môi trường	4	6645,42	1661,36	<0,0001	2,22	30,71	
	Tương tác	988	7256,28	7,34	<0,0001	2,38	33,54	
	Lỗi	2478	479,53	0,19	<0,0001			
	Tổng biến thể	3719	21636,40					

Ghi chú: # chiều dài hạt (SL), chiều rộng hạt (SW), độ dày hạt (ST) và trọng lượng 100 hạt (HSW); <sup>a</sup> Nguồn của biến thể; <sup>b</sup> Độ tự do; <sup>c</sup> Tổng của độ lệch bình phương; <sup>d</sup> Trung bình bình phương; <sup>e</sup> Giá trị P của F-test (kiểm tra giả thuyết chung); <sup>f</sup> Thành phần phương sai cho các nguồn biến thể; <sup>g</sup> Tỷ lệ của biến thể; <sup>h</sup> Di truyền theo nghĩa rộng; <sup>i</sup> CV; <sup>j</sup> Độ lệch; <sup>k</sup> Độ nhọn.

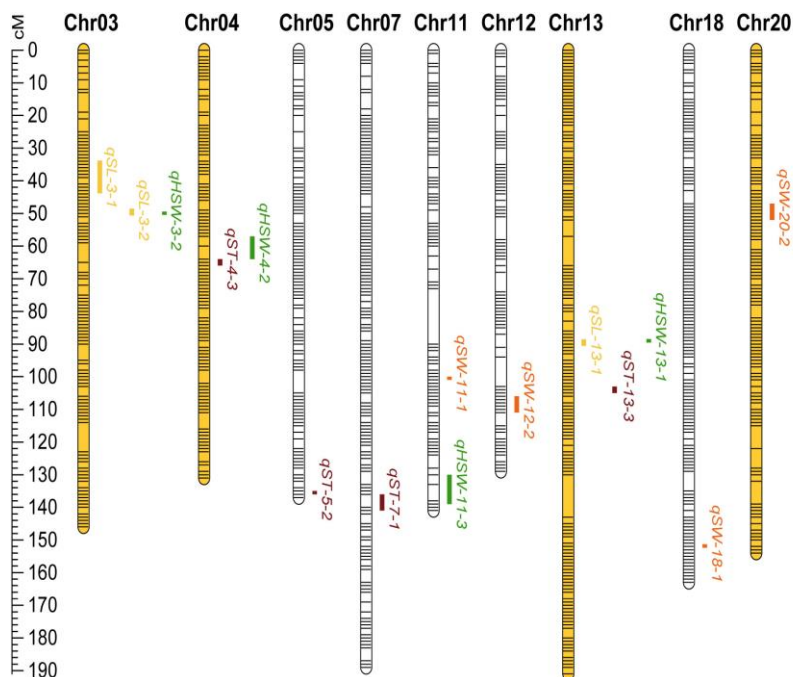
Xét về tỷ lệ của tổng biến thể chiếm sự biến đổi của các nguồn khác nhau (Bảng 2), nguồn biến thể lớn nhất trong SL và SW là kiểu gen, trong khi nguồn biến thể lớn nhất ở ST và HSW là tương tác kiểu gen × môi trường. Chịu ảnh hưởng lớn nhất của môi trường là HSW, chiếm 30,71% tổng biến thể. Tiếp theo là SL, SW và ST, lần lượt chiếm 19,51%, 18,19% và 17,53% trong tổng số biến thể. Kết quả trên cho thấy tính trạng trọng lượng hạt chịu ảnh hưởng nhiều hơn của môi trường so với tính trạng kích thước hạt.

### 3.4. Xác định QTL ổn định cho các tính trạng kích thước và trọng lượng hạt

Tổng số 85 QTL được phát hiện về kích thước và trọng lượng hạt trong 6 môi trường có PVE nằm trong khoảng 3,14–21,93%, trong đó 19 QTL được xác định cho SL, 23 cho



SW, 18 cho ST và 25 cho HSW (Bổ sung Bảng 1–4). Bằng cách đối chiếu và phân tích, tổng số 18 QTL ổn định đã được xác định (Bảng 3), bao gồm 3 cho SL, 6 cho SW, 4 cho ST và 5 cho HSW (Hình 4). Một QTL, *qSL-3-1*, liên quan đến SL đã được phát hiện trong tất cả các môi trường và nó giải thích hơn 10% PV trong ba môi trường trong khoảng 11,93–15,91%. Ba QTL (*qHSW-11-3*, *qHSW-20-3* và *qSW-20-3*) đã được phát hiện trong bốn môi trường. Trong số đó, *qSW-20-3* giải thích lần lượt 9,22%, 10,79%, 20,89% và 21,93% PV trong bốn môi trường. 14 QTL khác phân bố trên các nhiễm sắc thể 03, 04, 05, 07, 11, 12, 13, 18 và 20 được phát hiện trong ba môi trường. Tóm lại, việc tích hợp PVE cho QTL, số lần phát hiện trong các môi trường khác nhau và các tài liệu tham khảo có liên quan, *qSL-3-1* và *qSW-20-3* có thể được coi là QTL ổn định mới với các tác dụng chính. Trong khi đó, các QTL ổn định ở trên là tài liệu tham khảo quan trọng cho việc khai quật các gen kiểm soát các tính trạng về kích thước và trọng lượng của hạt.



**Hình 4.** Bản đồ liên kết của các QTL ổn định đối với các tính trạng kích thước và trọng lượng hạt đậu tương. Thuộc ở phía bên trái cho biết khoảng cách giữa các marker sử dụng cM (centimorgan) làm đơn vị. Các thanh có điểm nhấn màu vàng biểu thị nhiễm sắc thể chứa các nhóm QTL. Giữa các thanh trong biểu đồ, màu vàng biểu thị QTL cho SL; màu cam biểu thị QTL cho SW; màu nâu đỏ biểu thị QTL cho ST; màu xanh biểu thị QTL cho HSW.

### 3.5. Xác định các nhóm QTL

Theo tiêu chí phân loại của các nhóm QTL, chúng tôi đã thu được năm nhóm QTL (Bảng 4) từ kết quả bản đồ QTL (Bảng bổ sung 1–4). Năm nhóm QTL này được phân bố trên các nhiễm sắc thể 03, 04, 13 và 20. Tất cả chúng đều được liên kết với HSW, điều này đã xác minh rằng kích thước hạt có liên quan chặt chẽ với HSW từ một khía cạnh khác. Bốn nhóm QTL liên quan đến SL là *qLH3-1*, *qLH3-2*, *qLWTH4* và *qLH13*; hai nhóm QTL liên quan đến SW là *qLWTP4* và *qWTH20*; hai nhóm QTL liên quan đến ST là *qLWTP4* và *qWTP20*. Xét về số lượng tính trạng được kiểm soát, một nhóm QTL cho 4 tính trạng là *qLWTP4*, một nhóm QTL cho 3 tính trạng là *qWTH20* và 3 nhóm QTL cho 2 tính trạng là *qLH3-1*, *qLH3-2* và *qLH13*. Trong số đó, *qLP3-1* và *qLP3-2* chứa (các) QTL chính và ổn định nằm trên nhánh trước của nhiễm sắc thể 03 ở các vị trí vật lý tương ứng trong khoảng từ 4387438–6456915bp và từ 8784766–17199876bp. Tương tự, *qWTH20* cũng chứa các QTL chính và ổn định và nằm ở vị trí vật lý của nhiễm sắc thể 20 giữa 12056958 và 33222868bp.

**Bảng 3.** Các QTL ổn định liên quan đến các tính trạng kích thước và trọng lượng hạt đối với quần thể RIL GB13 trên các môi trường khác nhau.

QTL <sup>#</sup>	Chr <sup>a</sup>	Khoảng thời gian	Vị trí (cM)	PVE(%) <sup>h</sup>						LOD <sup>i</sup>						ADD <sup>j</sup>					
				E1 <sup>b</sup>	E2 <sup>c</sup>	E3 <sup>d</sup>	E4 <sup>e</sup>	E5 <sup>f</sup>	E6 <sup>g</sup>	E1	E2	E3	E4	E5	E6	E1	E2	E3	E4	E5	E6
<i>qSL-3-1</i>	3	bin32-bin43	43,1	13,57	6,03	11,93	4,38	6,61	15,91	9,37	4,03	7,49	3,15	4,61	11,45	0,15	0,09	0,11	0,08	0,09	0,11
<i>qSL-3-2</i>	3	bin51-bin55	49,4			12,15		9,07	14,56			7,63		6,42	10,38			0,11		0,11	0,11
<i>qHSW-3-2</i>	3	bin53-bin55	50,1			5,12		4,41	6,81			3,68		3,20	5,43			0,38		0,37	0,37
<i>qHSW-4-2</i>	4	bin75-bin78	57,5	5,34		4,68			5,28	3,75		3,25			4,18	0,59		0,37			0,33
<i>qST-4-3</i>	4	bin77-bin83	65,2	4,86	8,61				6,10	3,18	5,44				4,70	0,06	0,08				0,04
<i>qST-5-2</i>	5	bin155-bin158	135,1	4,67			4,52		3,69	3,06			3,13		2,89	-0,06			-0,06		-0,04
<i>qST-7-1</i>	7	bin159-bin161	139,9				3,85	5,75	6,99				2,69	4,06	5,37				0,05	0,06	0,05
<i>qSW-11-1</i>	11	bin95-bin100	101	3,89			7,48		6,34	2,90			5,23		5,14	-0,06			-0,08		-0,05
<i>qHSW-11-3</i>	11	bin131-bin134	133,1		4,90		8,26	3,33	4,41		2,99		5,76	2,56	3,41		0,42		0,59	0,32	0,31
<i>qSW-12-2</i>	12	bin103-bin109	111,1			5,79		3,69	5,39			4,37		2,83	4,36			0,06		0,05	0,05
<i>qHSW-13-1</i>	13	bin115-bin116	90,0		4,34			5,19	7,21		2,91			3,79	5,69		0,39			0,40	0,38
<i>qSL-13-1</i>	13	bin115-bin117	90	3,47				4,12	4,47	2,55				3,03	3,54	0,08				0,07	0,06
<i>qST-13-3</i>	13	bin143-bin149	104,6		5,09	5,22			4,56		3,27	3,74			3,50		0,06	0,06			0,04
<i>qSW-18-1</i>	18	bin237-bin239	152,4	5,62		4,91			3,91	4,03		3,72			3,24	0,07		0,05			0,04
<i>qSW-20-2</i>	20	bin46-bin49	49,1	7,46	11,80	8,64				5,30	8,29	6,24				-0,08	-0,10	-0,07			
<i>qSW-20-3</i>	20	bin64-bin68	55,3	9,22		10,79		21,93	20,89	6,63		7,89		15,04	15,24	-0,09	-0,10	-0,08		-0,14	-0,10
<i>qHSW-20-3</i>	20	bin68-bin73	55,9		6,12	3,31		4,86	7,21		4,06	2,52		3,42	5,70		-0,47	-0,31		-0,38	-0,38
<i>qSW-20-4</i>	20	bin84-bin86	64,2			3,80		7,30	5,81			2,65		4,56	3,86			-0,05		-0,09	-0,05

Ghi chú: <sup>#</sup> Tên của từng QTL; <sup>a</sup> Nhiễm sắc thể; <sup>b</sup> Tại Zengcheng năm 2018; <sup>c</sup> Tại Guangzhou năm 2019; <sup>d</sup> Tại Zengcheng năm 2019; <sup>e</sup> Tại Guangzhou năm 2020; <sup>g</sup> Môi trường kết hợp; <sup>h</sup> Biểu thể kiểu hình được giải thích; <sup>i</sup> Logarit của giá trị l; <sup>j</sup> Hiệu ứng phụ.

**Bảng 4.** Các nhóm QTL liên quan đến ít nhất hai tính trạng.

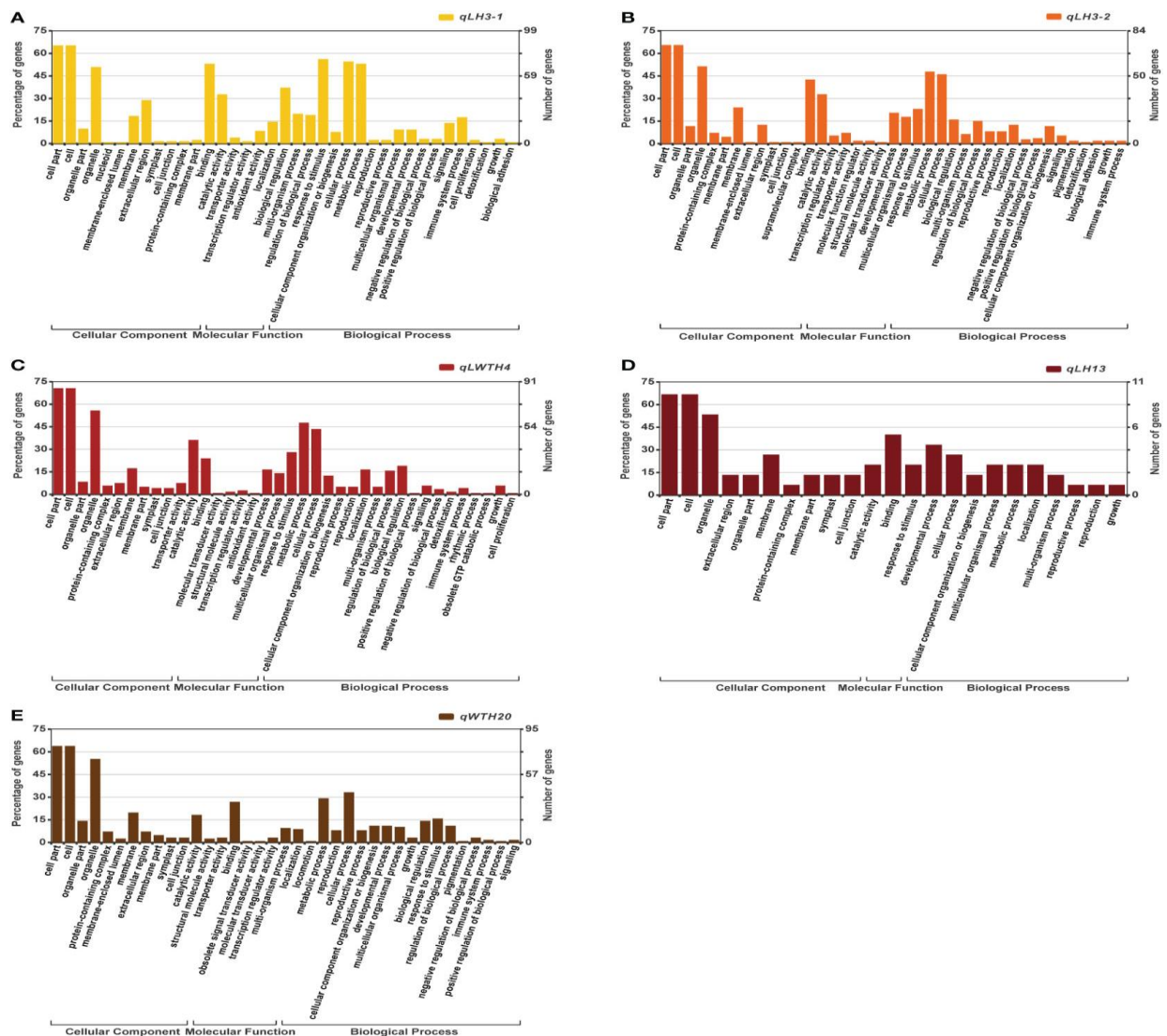
Nhóm	Chr <sup>a</sup>	Khoảng thời gian	Vị trí (bp)	QTLs <sup>b</sup>	ADD <sup>c</sup>	PVE(%) <sup>d</sup>	Env. <sup>e</sup>
<i>qLH3-1</i>	3	bin32- bin43	4387438- 6456915	Qsl-3-1	0,15	13,57	18ZC
					0,09	6,03	19GZ
					0,11	11,93	19ZC
					0,08	4,38	20GZ
					0,09	6,61	20ZC
				qHSW-3-1	0,11	15,91	CE
					0,33	3,77	19ZC
<i>qLH3-2</i>	3	bin51- bin55	8784766- 17199876	qSL-3-2	0,11	12,15	19ZC
					0,11	9,07	20ZC
					0,11	14,56	CE
				qHSW-3-2	0,38	5,12	19ZC
<i>qLWTH4</i>	4	bin74- bin83	9040807- 12316836		0,37	4,41	20ZC
					0,37	6,81	CE
				qSL-4-2	0,06	4,36	CE
				qSW-4-2	0,06	3,78	18ZC
					0,04	3,7	CE
				qST-4-3	0,06	4,86	18ZC
					0,08	8,61	19GZ
					0,04	6,1	CE
<i>qLH13</i>	13	bin115- bin117	27473391- 27671175	qHSW-4-2	0,59	5,34	18ZC
					0,37	4,68	19ZC
					0,33	5,28	CE
				qSL-13-1	0,08	3,47	18ZC
					0,07	4,12	20ZC
					0,06	4,47	CE
<i>qWTH20</i>	20	bin64- bin73	27890104- 33222868	qHSW-13-1	0,39	4,34	19GZ
					0,4	5,19	20ZC
					0,38	7,21	CE
				qSW-20-3	-0,09	9,22	18ZC
					-0,08	10,79	19ZC
					-0,14	21,93	20ZC
					-0,1	20,89	CE
				qST-20-2	-0,05	7,55	CE
				qHSW-20-3	-0,47	6,12	19GZ
	-0,31	3,31	19ZC				
	-0,38	4,68	20ZC				
	-0,38	7,21	CE				

Ghi chú: <sup>a</sup> Nhiễm sắc thể; <sup>b</sup> Các QTL riêng lẻ của nhóm; <sup>c</sup> Hiệu ứng phụ; <sup>d</sup> Giải thích biến thiên kiểu hình; <sup>e</sup> Môi trường

Tóm lại, *qLP3-1*, *qLP3-2* và *qWTH20* có thể là các nhóm QTL chứa các gen tác động chính quy định các tính trạng kích thước và trọng lượng hạt đậu tương; *qLWTP4* và *qLP13* có thể là các nhóm QTL chứa các gen tác động vi mô làm hài hòa việc kiểm soát các tính trạng về kích thước và trọng lượng của hạt đậu tương. Do đó, năm nhóm QTL này có thể cung cấp tài liệu tham khảo để khai thác các gen mục tiêu kiểm soát các tính trạng về kích thước và trọng lượng hạt và được sử dụng làm khoảng thời gian mục tiêu để phân tích làm giàu gen GO.

### 3.6. Phân tích làm giàu Gene GO

Bằng phân tích làm giàu Gene GO, chúng tôi thấy rằng hầu hết các gen trong năm nhóm QTL này đều tham gia vào các quá trình tế bào và quá trình trao đổi chất (Hình 5). Hầu hết các gen trong *qLP3-1* tham gia vào các quá trình trao đổi chất nội bào, điều hòa sinh học và phản ứng với các kích thích, và chủ yếu tham gia vào các hoạt động liên kết và xúc tác về mặt chức năng phân tử. Trong khi hầu hết các gen trong *qLP3-2* có chức năng tương tự như trong *qLP3-1*. Để so sánh, hầu hết các gen trong *qLWTP4* cũng tham gia vào các quá trình trao đổi chất nội bào, nhưng một số lượng nhỏ gen tham gia vào việc điều chỉnh sự tăng trưởng và tăng sinh tế bào. *qLP13* chỉ chứa 11 gen được chú thích, hầu hết trong số đó có liên quan đến quá trình tăng trưởng và tế bào. Tương tự, hầu hết các gen trong *qWTP20-2* đều tham gia vào quá trình trao đổi chất nội bào.



**Hình 5.** GO phân tích làm giàu thuật ngữ của các gen nằm trong năm nhóm QTL: (A) *qLH3-1*; (B) *qLH3-2*; (C) *qLWTH4*; (D) *qLH13*; (E) *qWTH20*.

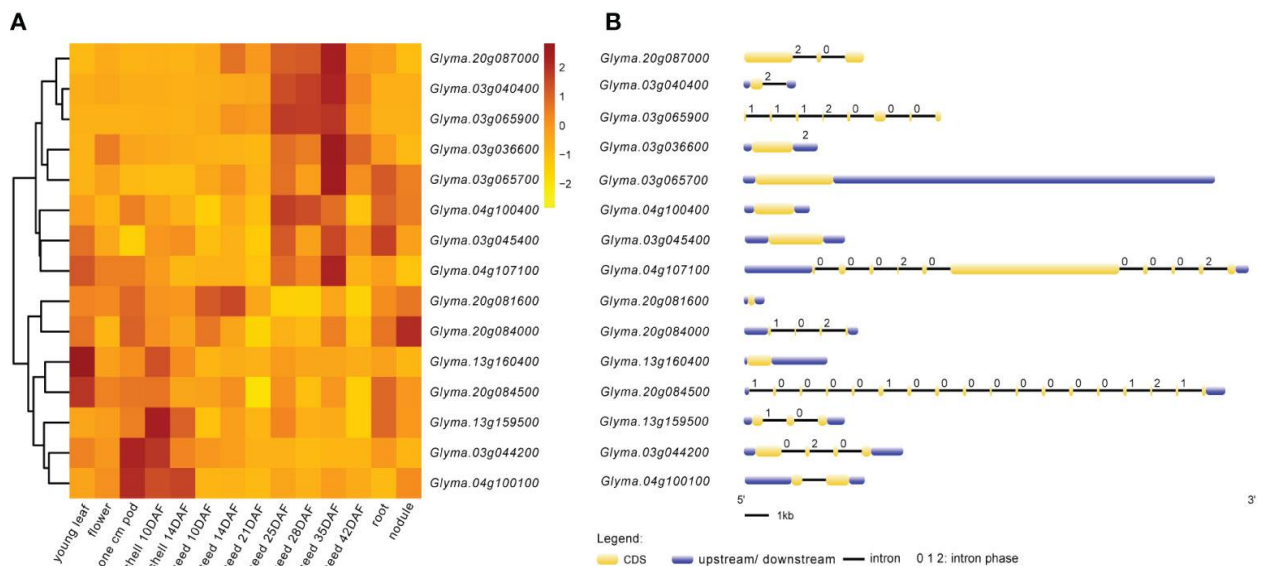
### 3.7. Các gen ứng cử viên trong nhóm QTL

Thông qua phân tích làm giàu gen GO ở trên, tiếp theo là sàng lọc biểu hiện gen, chú thích chức năng gen và phân tích biến thể trình tự, chúng tôi đã thu được tổng cộng 15 gen ứng cử viên (Bảng 5) có thể tham gia vào việc điều chỉnh các tính trạng kích thước và trọng lượng hạt. Trong đó, 6 gen nằm trên nhiễm sắc thể 03, 3 gen trên nhiễm sắc thể 04, 2 gen trên nhiễm sắc thể 13 và 4 gen trên nhiễm sắc thể 20. Trong số 15 gen được sàng lọc, *Glyma.03g040400* là gen có cùng chức năng với gen liên kết (*LOC\_Os07g19000*) với kích thước hạt gạo; *Glyma.03g045400* và *Glyma.20g084000* có liên quan đến quá trình methyl hóa DNA; *Glyma.04g100400*, *Glyma.20g081600* và *Glyma.20g087000* có liên quan đến hormone và có thể liên quan đến việc điều chỉnh các tính trạng kích thước và trọng lượng hạt bằng con đường truyền tín hiệu hormone; *Glyma.04g100100*, *Glyma.04g107100*, *Glyma.13g159500* và *Glyma.20g084500* có thể điều chỉnh tiêu cực kích thước và trọng lượng hạt thông qua con đường ubiquitin-proteasome. *Glyma.03g036600* và *Glyma.03g065700* được liên kết với quy định về yếu tố phiên mã và có thể tham gia vào quy định về kích thước hạt và các tính trạng trọng lượng thông qua con đường điều hòa yếu tố phiên mã; *Glyma.03g044200*, *Glyma.03g065900* và *Glyma.13g160400* có thể liên quan đến quá trình tổng hợp trao đổi chất của protein và dầu, đồng thời gián tiếp điều chỉnh các tính trạng kích thước và trọng lượng của hạt.

**Bảng 5.** Các gen ứng cử viên được xác định trong năm nhóm QTL được biểu hiện cao hoặc biểu hiện cụ thể trong hạt đậu tương.

Gen	Vị trí (bp)	Chú thích chức năng gen
<i>Glyma.03g036600</i>	4432910-4435002	Quy định tiêu cực của phiên mã; phát triển cơ quan; vận chuyển lipid gắn kết balmodulin
<i>Glyma.03g040400</i>	5050463-5052307	Ức chế protease (họ LTP); lưu trữ hạt giống
<i>Glyma.03g044200</i>	5586338-5589780	Hoạt động vận chuyển Protein; tích hợp với màng
<i>Glyma.03g045400</i>	5760642-5763461	Quá trình trao đổi chất; hoạt động methyltransferase
<i>Glyma.03g065700</i>	11896538-11909845	Điều hòa phiên mã; hoạt động của yếu tố phiên mã liên kết DNA theo trình tự cụ thể; phụ thuộc DNA
<i>Glyma.03g065900</i>	12064850-12067709	Quá trình trao đổi chất; hoạt động thủy phân; Hoạt động xúc tác
<i>Glyma.04g100100</i>	9177207-9180239	Quy định tiêu cực của biểu hiện gen; phản ứng với kích thích axit abscisic; hoạt động ligase ubiquitin-protein
<i>Glyma.04g100400</i>	9219564-9221412	Quá trình sinh tổng hợp sterol; phản ứng với kích thích oleuropein lactone
<i>Glyma.04g107100</i>	11239281-11250948	Điều hòa sinh trưởng mô phân sinh; Dây chằng Ubiquitin có liên quan đến sự thoái hóa protein synap
<i>Glyma.13g159500</i>	27476780-27479031	Các thành phần hệ thống Ubiquitin gợi ý về protein; tăng trưởng tế bào; hình thái tế bào
<i>Glyma.13g160400</i>	27606020-27608359	Vận chuyển lipid; liên kết lipid; protein kỵ nước

<i>Glyma.20g081600</i>	30778586-30779164	Điều hòa biểu hiện gen; phản ứng với kích thích auxin; liên kết protein
<i>Glyma.20g084000</i>	31536334-31538924	Methyl hóa ARN; xử lý RNA
<i>Glyma.20g084500</i>	31666190-31674959	Hoạt động ligase Ubiquitin-protein; phổ biến protein; Phức hợp ligase CUL4-RING ubiquitin
<i>Glyma.20g087000</i>	32527435-32530216	Quy định tiêu cực của con đường truyền tín hiệu qua trung gian ethylene; quy định phiên mã; liên kết ethylene



**Hình 6.** Các kiểu biểu hiện và cấu trúc của các gen ứng cử viên. (A) Bản đồ nhiệt trực quan của các gen ứng cử viên được biểu hiện trong các mô và thời kỳ khác nhau của cây đậu tương. Dữ liệu biểu hiện của 15 gen ứng cử viên trong nghiên cứu này đã được tải xuống từ SoyBase (<http://www.soybase.org/soyseq/>). Chuẩn hóa dữ liệu biểu hiện bởi các gen khác nhau. (B) Sơ đồ cấu trúc gen của các gen ứng cử viên. Màu vàng biểu thị CDS (trình tự DNA mã hóa); màu xanh biểu thị ngược dòng hoặc xuôi dòng của CDS chứa 5'UTR và 3'UTR. Số Ả Rập chỉ số lượng pha intron.

**Bảng 6.** Các biến thể trình tự giữa các dòng bố mẹ của các gen ứng cử viên.

Gen	Chr <sup>a</sup>	Vị trí (bp) <sup>b</sup>	Guizao1 <sup>c</sup>	B13 <sup>d</sup>	Region <sup>e</sup>
<i>Glyma.03g036600</i>	3	4433324	T	TA	UTR3
<i>Glyma.03g040400</i>	3	5051148	A	ACG	intronic
	3	5051173	A	ATCACC	intronic
	3	5051386	T	TC	intronic
	3	5052301	G	GAT	UTR3
<i>Glyma.03g044200</i>	3	5588615	TCA	T	intronic
	3	5588681	AAC	A	intronic
	3	5588699	AAC	A	intronic
<i>Glyma.03g045400</i>	3	5760781	T	TTC	UTR3
	3	5760989	TG	T	UTR3
	3	5761009	A	AAAACAAGAG	UTR3

	3	5761114	AACTCTC	A	UTR3
	3	5763119	T	TATG	UTR5
<i>Glyma.03g065700</i>	3	11907032	G	GA	UTR3
<i>Glyma.03g065900</i>	3	NA	NA	NA	NA
<i>Glyma.04g100100</i>	4	9179539	GA	GAAAACAAAAC	UTR5
<i>Glyma.04g100400</i>	4	NA	NA	NA	NA
<i>Glyma.04g107100</i>	4	11246102	TA	T	intronic
<i>Glyma.13g159500</i>	13	NA	NA	NA	NA
<i>Glyma.13g160400</i>	13	27607021	TA	T	UTR3
<i>Glyma.20g081600</i>	20	NA	NA	NA	NA
<i>Glyma.20g084000</i>	20	31536882	A	AAAAACCCTACATCTTCAT	intronic
	20	31537182	ATT	A	intronic
	20	31537333	TTA	T	intronic
	20	31537718	A	AGAATCAATGTCATT	intronic
	20	31538735	TTA	T	UTR3
<i>Glyma.20g084500</i>	20	31666328	GTCA	G	UTR3
	20	31669438	A	AAC	intronic
	20	31673766	TCG	T	intronic
	20	31674346	T	TAA	intronic
	20	31674567	T	TA	intronic
<i>Glyma.20g084500</i>	20	NA	NA	NA	NA

#### 4. THẢO LUẬN

Dựa trên kết quả phân tích tương quan giữa kích thước hạt và trọng lượng hạt trong nghiên cứu này,  $R^2$  giữa chiều rộng hạt và trọng lượng hạt lớn nhất trong hầu hết các môi trường, vì vậy chúng tôi đề xuất có thể chọn các giống có hạt rộng để cải thiện cho tính trạng trọng lượng hạt đậu tương. Tuy nhiên, từ kết quả của các QTL được nội địa hóa cho các tính trạng kích thước và trọng lượng hạt đậu tương, hầu hết các QTL cho HSW trùng nhau hoặc gần với khoảng QTL cho SL. Nó gợi ý rằng hầu hết các gen có đa hiệu ứng có thể là các gen phối hợp kiểm soát trên cả SL và HSW. Tuy nhiên, cơ chế quản lý cụ thể vẫn chưa rõ ràng (He và cs, 2021) và cần được nghiên cứu thêm.

Trong nghiên cứu này, tổng số 60 QTL liên quan đến tính trạng kích thước hạt và 25 QTL liên quan đến tính trạng trọng lượng hạt đã được xác định trong quần thể RIL GB13 trong năm môi trường tự nhiên và môi trường kết hợp của chúng. Mười chín QTL cho SL bao gồm 4 QTL ổn định và 15 QTL nhạy cảm (Bảng bổ sung 1); 23 QTL cho chiều rộng hạt bao gồm 6 QTL ổn định và 17 QTL nhạy cảm (Bảng bổ sung 2); 18 QTL cho ST bao gồm 5 QTL ổn định và 13 QTL nhạy cảm (Bảng bổ sung 3). 25 QTL liên quan đến HSW bao gồm 5 QTL ổn định và 20 QTL nhạy cảm (Bảng bổ sung 4). Bằng cách so sánh kết quả với những nghiên cứu trước đây, mười QTL nhạy cảm với SL (*qSL-2-1*, *qSL-3-3*, *qSL-4-1*, *qSL-4-2*, *qSL-4-3*, *qSL-6-2*, *qSL-11-1*, *qSL-11-2*, *qSL-18-2* và *qSL-20-1*) trùng lặp với khoảng thời gian của các kết quả nội địa hóa trước đó hoặc nằm trong các tập hợp con của chúng (Salas và cs, 2006; Xu và cs, 2011; Niu và cs, 2013; Yang và cs, 2017;

Cui và cs, 2020; Hina và cs, 2020; Kumawat và Xu, 2021). Tám QTL nhạy cảm cho SW (*qSW-2-1*, *qSW-2-2*, *qSW-4-1*, *qSW-9-1*, *qSW-10-1*, *qSW-15-1*, *qSW-20-1*, và *qSW-20-5*) trùng với hoặc nằm trong một tập hợp con của các khoảng kết quả nội địa hóa trước đó (Xu và cs, 2011; Niu và cs, 2013; Yang và cs, 2017; Hina và cs, 2020). Năm QTL nhạy cảm cho ST (*qST-4-4*, *qST-5-2*, *qST-6-1*, *qST-13-1* và *qST-20-3*) trùng lặp hoặc bao gồm các khoảng của kết quả nội địa hóa trước đó (Xu và cs, 2011; Niu và cs, 2013; Hina và cs, 2020). Năm QTL nhạy cảm cho HSW (*qHSW-4-3*, *qHSW-13-2*, *qHSW-17-2*, *qHSW-20-1* và *qHSW-20-5*) được phát hiện có vùng chồng lấp với kết quả nội địa hóa của nghiên cứu trước đây (Mansur và cs, 1993; Teng và cs, 2009; Han và cs, 2012; Kato và cs, 2014). Nhìn chung, kết quả của nghiên cứu hiện tại tương quan tốt với kết quả của các nghiên cứu trước đây.

Bên cạnh đó, so sánh kết quả nội địa hóa các tính trạng kích thước hạt trong nghiên cứu này với kết quả của các nghiên cứu nội địa hóa các tính trạng khác có liên quan, 8 QTL (*qSL-4-1*, *qSL-17-1*, *qSL-20-1*, *qSW-12-2*, *qSW-15-2*, *qSW-17-1*, *qST-13-3*, *qST-20-1*) đã được tìm thấy trùng lặp hoặc trùng lặp khoảng cách với các QTL liên quan đến tính trạng năng suất từ các nghiên cứu trước đó (Mansur và cs, 1993; Kabelka và cs, 2004; Du và cs, 2009; Palomeque và cs, 2009; Liu và cs, 2011; Wang X. và cs, 2014). Mười lăm QTL (*qSL-6-2*, *qSL-13-1*, *qSL-17-1*, *qSL-18-1*, *qSL-20-1*, *qSW-11-1*, *qSW-12-1*, *qSW-12-2*, *qSW-20-2*, *qSW-20-3*, *qST-5-2*, *qST-13-1*, *qST-13-3*, *qST-20-1*, *qST-20-2*) các QTL chồng chéo liên kết với các tính trạng liên quan đến protein, dầu và axit béo (Sebolt và cs, 2000; Tajuddin và cs, 2003; Bachlava và cs, 2009; Qi và cs, 2011; Eskandari và cs, 2013; Mao và cs, 2013; Wang J. và cs, 2014; Fan và cs, 2015; Warrington và cs, 2015; Han và cs, 2015; Leite và cs, 2016; Fliege và cs, 2022). Năm QTL (*qSW-15-1*, *qSW-15-3*, *qSW-18-2*, *qST-4-1* và *qSW-5-2*) có khoảng thời gian trùng nhau với các QTL liên quan đến đóng hạt, số quả và độ chín của đậu tương (Yao và cs, 2015; Ning và cs, 2018). Hai QTL (*qSL-5-1* và *qST-5-2*) có khoảng thời gian trùng lặp với các QTL liên quan đến isoflavonoid (Yang và cs, 2011; Cai và cs, 2018). *qSW-4-1* trùng lặp với các QTL liên quan đến giai đoạn cây con kéo dài (Yue và cs, 2017). Các nhà nghiên cứu đôi khi nghiên cứu trọng lượng hạt đậu tương để lập bản đồ QTL kết hợp với hàm lượng protein và dầu (Fasoula và cs, 2004; Pathan và cs, 2013; Yesudas và cs, 2013; Wang X. và cs, 2014; Whiting và cs, 2020; He và cs, 2021). Trong nghiên cứu này, *qHSW-13-1*, *qHSW-15-1*, *qHSW-17-2*, *qHSW-20-2* và QTL ổn định *qHSW-20-3* có các khoảng trùng lặp với các QTL liên quan đến hàm lượng protein hoặc dầu được nghiên cứu trước đó (Panthee và cs, 2005; Mao và cs, 2013; Wang X. và cs, 2014; Han và cs, 2015; Fliege và cs, 2022). Bên cạnh đó, *qHSW-15-1* và *qHSW-17-1* cũng có khoảng thời gian trùng lặp với QTL liên quan đến năng suất đã được nghiên cứu trước đó (Palomeque và cs, 2009; Liu và cs, 2011). Tóm lại, nhiều QTL có tác dụng chính ổn định liên quan đến tính trạng kích thước và trọng lượng hạt trong nghiên cứu này phù hợp với kết quả của các nghiên cứu trước đây và tương quan với các QTL liên quan đến các tính trạng khác bao gồm protein, hàm lượng dầu, số quả, năng suất, độ chín và giai đoạn cây con kéo dài. Do đó, chúng tôi cho rằng các gen quy định tính trạng kích thước và trọng lượng hạt đậu tương có thể tương quan với các gen quy định các tính trạng khác, đặc biệt là hàm lượng protein và dầu, vì vậy các gen quy định quá trình tổng hợp hoặc chuyển hóa hàm lượng protein và dầu cũng được sử dụng làm cơ sở cho việc suy đoán gen ứng cử viên trong nghiên cứu này.

Ngoài ra, so với đậu tương, lúa (*Oryza sativa* L.) đã được nghiên cứu chuyên sâu hơn nhiều so với đậu tương về kích thước và trọng lượng hạt. Hiện tại, các gen liên quan đến kích thước và trọng lượng hạt, chẳng hạn như *GW2* (Song và cs, 2007), *OsGRF4* (Li và



cs, 2016), *GS3* (Fan và cs, 2006), *qSW5* (Shomura và cs, 2008), *GW5* (Weng và cs, 2008), *qGL3* (Zhang và cs, 2012), *LGY3* (Liu Q. và cs, 2018), *GS5* (Li và Li, 2016), *GS6* (Sun và cs, 2013), *GLW7* (Wang và cs, 2015), v.v., đã được nhân bản. Hơn nữa, các con đường điều hòa của gen chủ yếu bao gồm con đường ubiquitin-proteasome, con đường truyền tín hiệu protein G, con đường truyền tín hiệu protein kinase được kích hoạt bằng Mitogen (MAPK), con đường truyền tín hiệu hormone và con đường điều hòa yếu tố phiên mã, v.v. (Li và Li, 2016). Trong số đó, trong con đường ubiquitin-proteasome, E3 ubiquitin ligase đóng vai trò chính và điều chỉnh tiêu cực chiều rộng hạt và trọng lượng hạt (Song và cs, 2007). Trong con đường truyền tín hiệu hormone, Brassinolide (BR) và Auxin (IAA) có thể thúc đẩy sự phát triển và mở rộng của tế bào (Li và cs, 2019), nó chủ yếu điều chỉnh sự phát triển của glume để điều chỉnh tích cực kích thước hạt lúa (Li và cs, 2011). Do đó, trong nghiên cứu này, các gen ứng cử viên đã được lựa chọn và dự đoán bằng cách sàng lọc các gen trong năm nhóm dựa trên biểu hiện cụ thể của chúng ở cây đậu tương, kết hợp với các chức năng tương đồng hoặc giống hệt nhau của các gen liên quan đến tính trạng kích thước hạt và trọng lượng ở cây lúa.

QTL chính và ổn định, *qSW-20-3*, được phát hiện trong bốn môi trường có PVE thay đổi từ 9,22 – 21,93%, được chứa trong nhóm *qWTP20-2*. Bằng cách phân tích sự biến đổi trình tự, chúng tôi thấy rằng trong số các gen ứng cử viên trong khoảng QTL này, chỉ có *Glyma.20g084000* và *Glyma.20g084500* là khác nhau về trình tự của các dòng bố mẹ. Chúng không được biểu hiện cụ thể ở giai đoạn trưởng thành của hạt ở đậu tương, hơn nữa, chúng có tác dụng phụ tiêu cực, nhưng chúng có liên quan đến quá trình methyl hóa RNA và phổ biến protein tương ứng. Do đó, chúng tôi suy đoán rằng các gen trong nhóm này có thể gián tiếp điều chỉnh kích thước và trọng lượng hạt bằng cách điều chỉnh các tính trạng khác. Trong khi một QTL chính và ổn định khác, *qSL-3-1*, được phát hiện trong sáu môi trường có PVE thay đổi từ 4,38 – 15,91%, được chứa trong nhóm *qLP3-1*. So sánh sự biến đổi trình tự giữa các dòng bố mẹ đối với các gen ứng cử viên trong nhóm này, chúng tôi thấy rằng năm gen trải qua sự biến đổi tự nhiên giữa các dòng bố mẹ và tất cả chúng đều được biểu hiện đặc biệt và cao ở giai đoạn trưởng thành của hạt ở đậu tương. Hơn nữa, chúng có tác dụng phụ tích cực. Do đó, chúng tôi suy đoán rằng năm gen ứng cử viên trong nhóm *qLP3-1* điều chỉnh tích cực kích thước và trọng lượng hạt thông qua các con đường nhất định dựa trên chú thích gen của chúng. Tuy nhiên, đường dẫn tín hiệu chính xác vẫn cần được nghiên cứu thêm.

Điều quan trọng là phải nghiên cứu mạng lưới điều hòa phân tử của các tính trạng liên quan đến năng suất ở cây đậu tương. Và các locus và gen ứng cử viên trong nghiên cứu này cung cấp cơ sở lý thuyết và nguồn gen quan trọng để giải quyết vấn đề tắc nghẽn về năng suất đậu tương, vấn đề này xứng đáng được nghiên cứu thêm ở cấp độ phân tử.

## TÀI LIỆU THAM KHẢO

Bachlava, E., Dewey, R. E., Burton, J. W., Cardinal, A. J. (2009). Mapping and comparison of quantitative trait loci for oleic acid seed content in two segregating soybean populations. *Crop Sci.* 49 (2), 433–442. doi: 10.2135/cropsci2008.06.0324

[CrossRef Full Text](#) | [Google Scholar](#)

Cai, Z., Cheng, Y., Ma, Z., Liu, X., Ma, Q., Xia, Q., et al. (2018). Fine-mapping of QTLs for individual and total isoflavone content in soybean (*Glycine max* L.) using a high-density genetic map. *Theor. Appl. Genet.* 131 (3), 555–568. doi: 10.1007/s00122-017-3018-x

[PubMed Abstract](#) | [CrossRef Full Text](#) | [Google Scholar](#)

- Chapman, A., Pantalone, V. R., Ustun, A., Allen, F. L., Landau-Ellis, D., Trigiano, R. N., et al. (2003). Quantitative trait loci for agronomic and seed quality traits in an F<sub>2</sub> and F<sub>4:6</sub> soybean population. *Euphytica*. 129 (3), 387–393. doi: 10.1023/A:1022282726117  
[CrossRef Full Text](#) | [Google Scholar](#)
- Csanadi, G., Vollmann, J., Stift, G., Lelley, T. (2001). Seed quality QTLs identified in a molecular map of early maturing soybean. *Theor. Appl. Genet.* 103 (6-7), 912–919. doi: 10.1007/s001220100621  
[CrossRef Full Text](#) | [Google Scholar](#)
- Cui, B., Chen, L., Yang, Y., Liao, H. (2020). Genetic analysis and map-based delimitation of a major locus qSS3 for seed size in soybean. *Plant Breeding*. 139 (6), 1145–1157. doi: 10.1111/pbr.12853  
[CrossRef Full Text](#) | [Google Scholar](#)
- Du, W., Wang, M., Fu, S., Yu, D. (2009). Mapping QTLs for seed yield and drought susceptibility index in soybean (*Glycine max* l.) across different environments. *J. Genet. Genomics* 36 (12), 721–731. doi: 10.1016/S1673-8527(08)60165-4  
[PubMed Abstract](#) | [CrossRef Full Text](#) | [Google Scholar](#)
- Elattar, M. A., Karikari, B., Li, S., Song, S., Cao, Y., Aslam, M., et al. (2021). Identification and validation of major QTLs, epistatic interactions, and candidate genes for soybean seed shape and weight using two related RIL populations. *Front. Genet.* 12. doi: 10.3389/fgene.2021.666440  
[CrossRef Full Text](#) | [Google Scholar](#)
- Eskandari, M., Cober, E. R., Rajcan, I. (2013). Genetic control of soybean seed oil: I. QTL and genes associated with seed oil concentration in RIL populations derived from crossing moderately high-oil parents. *Theor. Appl. Genet.* 126 (2), 483–495. doi: 10.1007/s00122-012-1995-3  
[PubMed Abstract](#) | [CrossRef Full Text](#) | [Google Scholar](#)
- Fan, S., Li, B., Yu, F., Han, F., Yan, S., Wang, L., et al. (2015). Analysis of additive and epistatic quantitative trait loci underlying fatty acid concentrations in soybean seeds across multiple environments. *Euphytica*. 206 (3), 689–700. doi: 10.1007/s10681-015-1491-3  
[CrossRef Full Text](#) | [Google Scholar](#)
- Fan, C., Xing, Y., Mao, H., Lu, T., Han, B., Xu, C., et al. (2006). GS3, a major QTL for grain length and weight and minor QTL for grain width and thickness in rice, encodes a putative transmembrane protein. *Theor. Appl. Genet.* 112 (6), 1164–1171. doi: 10.1007/s00122-006-0218-1  
[PubMed Abstract](#) | [CrossRef Full Text](#) | [Google Scholar](#)
- Fasoula, V. A., Harris, D. K., Boerma, H. R. (2004). Validation and designation of quantitative trait loci for seed protein, seed oil, and seed weight from two soybean populations. *Crop Sci.* 44 (4), 1218–1225. doi: 10.2135/cropsci2004.1218  
[CrossRef Full Text](#) | [Google Scholar](#)
- Fliege, C. E., Ward, R. A., Vogel, P., Nguyen, H., Quach, T., Guo, M., et al. (2022). Fine mapping and cloning of the major seed protein quantitative trait loci on soybean chromosome 20. *Plant J.* 110 (1), 114–128. doi: 10.1111/tpj.15658  
[PubMed Abstract](#) | [CrossRef Full Text](#) | [Google Scholar](#)
- Gao, H. H., Gu, Y., Jiang, H. W., Li, Y. Y., Xin, D. W., Liu, C. Y., et al. (2019). Mapping and analysis of QTLs related to seed length and seed width in *Glycine max*. *Plant Breeding*. 138 (6), 733–740. doi: 10.1111/pbr.12745  
[CrossRef Full Text](#) | [Google Scholar](#)

- Han, Y., Teng, W., Wang, Y., Zhao, X., Wu, L., Li, D., et al. (2015). Unconditional and conditional QTL underlying the genetic interrelationships between soybean seed isoflavone, and protein or oil contents. *Plant Breeding*. 134 (3), 300–309. doi: 10.1111/pbr.12259  
[CrossRef Full Text](#) | [Google Scholar](#)
- Han, Y., Xie, D., Teng, W., Sun, J., Li, W. (2012). QTL underlying developmental behaviour of 100-seed weight of soybean. *Plant Breeding*. 131 (5), 600–606. doi: 10.1111/j.1439-0523.2012.01987.x  
[CrossRef Full Text](#) | [Google Scholar](#)
- He, Q., Xiang, S., Yang, H., Wang, W., Shu, Y., Li, Z., et al. (2021). A genome-wide association study of seed size, protein content, and oil content using a natural population of sichuan and chongqing soybean. *Euphytica*. 217 (11), 198. doi: 10.1007/s10681-021-02931-8  
[CrossRef Full Text](#) | [Google Scholar](#)
- Hina, A., Cao, Y., Song, S., Li, S., Sharmin, R. A., Elattar, M. A., et al. (2020). High-resolution mapping in two RIL populations refines major QTL “hotspot” regions for seed size and shape in soybean (*Glycine max* l.). *Int. J. Mol. Sci.* 21 (3), 1040. doi: 10.3390/ijms21031040  
[PubMed Abstract](#) | [CrossRef Full Text](#) | [Google Scholar](#)
- Hu, Z., Zhang, H., Kan, G., Ma, D., Zhang, D., Shi, G., et al. (2013). Determination of the genetic architecture of seed size and shape *via* linkage and association analysis in soybean (*Glycine max* l. merr.). *Genetica*. 141 (4-6), 247–254. doi: 10.1007/s10709-013-9723-8  
[PubMed Abstract](#) | [CrossRef Full Text](#) | [Google Scholar](#)
- Hyten, D. L., Pantalone, V. R., Sams, C. E., Saxton, A. M., Landau-Ellis, D., Stefaniak, T. R., et al. (2004). Seed quality QTL in a prominent soybean population. *Theor. Appl. Genet.* 109 (3), 552–561. doi: 10.1007/s00122-004-1661-5  
[PubMed Abstract](#) | [CrossRef Full Text](#) | [Google Scholar](#)
- Jiang, B. Z., Li, M., Cheng, Y. B., Cai, Z. D., Ma, Q. B., Jiang, Z., et al. (2019). Genetic mapping of powdery mildew resistance genes in soybean by high-throughput genome-wide sequencing. *Theor. Appl. Genet.* 132 (6), 1833–1845. doi: 10.1007/s00122-019-03319-y  
[PubMed Abstract](#) | [CrossRef Full Text](#) | [Google Scholar](#)
- Jia, J., Wang, H., Cai, Z., Wei, R., Huang, J., Xia, Q., et al. (2022). Identification and validation of stable and novel quantitative trait loci for pod shattering in soybean [*Glycine max* (L.) merr.]. *J. Integr. Agr.* 21 (11), 3169–3184. doi: 10.1016/j.jia.2022.08.082  
[CrossRef Full Text](#) | [Google Scholar](#)
- Kabelka, E. A., Diers, B. W., Fehr, W. R., LeRoy, A. R., Baianu, I. C., You, T., et al. (2004). Putative alleles for increased yield from soybean plant introductions. *Crop Sci.* 44 (3), 784. doi: 10.2135/cropsci2004.0784  
[CrossRef Full Text](#) | [Google Scholar](#)
- Kato, S., Sayama, T., Fujii, K., Yumoto, S., Kono, Y., Hwang, T., et al. (2014). A major and stable QTL associated with seed weight in soybean across multiple environments and genetic backgrounds. *Theor. Appl. Genet.* 127 (6), 1365–1374. doi: 10.1007/s00122-014-2304-0  
[PubMed Abstract](#) | [CrossRef Full Text](#) | [Google Scholar](#)
- Kulkarni, K. P., Asekova, S., Lee, D., Bilyeu, K., Song, J. T., Lee, J. (2017). Mapping QTLs for 100-seed weight in an interspecific soybean cross of williams 82 (*Glycine max*) and PI 366121 (*Glycine soja*). *Crop Pasture Sci.* 68 (2), 148–155. doi: 10.1071/CP16246  
[CrossRef Full Text](#) | [Google Scholar](#)
- Kumawat, G., Xu, D. (2021). A major and stable quantitative trait locus qSS2 for seed size and shape traits in a soybean RIL population. *Front. Genet.* 12. doi: 10.3389/fgene.2021.646102  
[PubMed Abstract](#) | [CrossRef Full Text](#) | [Google Scholar](#)

- Lee, S. H., Park, K. Y., Lee, H. S., Park, E. H., Boerma, H. R. (2001). Genetic mapping of QTLs conditioning soybean sprout yield and quality. *Theor. Appl. Genet.* 103 (5), 702–709.  
doi: 10.1007/s001220100595  
[CrossRef Full Text](#) | [Google Scholar](#)
- Leite, D. C., Pinheiro, J. B., Campos, J. B., Di Mauro, A. O., Unêda-Trevisoli, S. H. (2016). QTL mapping of soybean oil content for marker-assisted selection in plant breeding program. *Genet. Mol. Res.* 15 (1), grm7685. doi: 10.4238/gmr.15017685  
[CrossRef Full Text](#) | [Google Scholar](#)
- Li, Y., Fan, C., Xing, Y., Jiang, Y., Luo, L., Sun, L., et al. (2011). Natural variation in GS5 plays an important role in regulating grain size and yield in rice. *Nat. Genet.* 43 (12), 1266–1269.  
doi: 10.1038/ng.977  
[PubMed Abstract](#) | [CrossRef Full Text](#) | [Google Scholar](#)
- Li, S., Gao, F., Xie, K., Zeng, X., Cao, Y., Zeng, J., et al. (2016). The OsmiR396c-OsGRF4-OsGIF1 regulatory module determines grain size and yield in rice. *Plant Biotechnol. J.* 14 (11), 2134–2146. doi: 10.1111/pbi.12569  
[PubMed Abstract](#) | [CrossRef Full Text](#) | [Google Scholar](#)
- Li, N., Li, Y. (2016). Signaling pathways of seed size control in plants. *Curr. Opin. Plant Biol.* 33, 23–32. doi: 10.1016/j.pbi.2016.05.008  
[PubMed Abstract](#) | [CrossRef Full Text](#) | [Google Scholar](#)
- Li, M., Liu, Y., Wang, C., Yang, X., Li, D., Zhang, X., et al. (2020). Identification of traits contributing to high and stable yields in different soybean varieties across three chinese latitudes. *Front. Plant Sci.* 10, 1642. doi: 10.3389/fpls.2019.01642  
[PubMed Abstract](#) | [CrossRef Full Text](#) | [Google Scholar](#)
- Liu, Q., Han, R., Wu, K., Zhang, J., Ye, Y., Wang, S., et al. (2018). G-Protein  $\beta\gamma$  subunits determine grain size through interaction with MADS-domain transcription factors in rice. *Nat. Commun.* 9 (1), 852. doi: 10.1038/s41467-018-03047-9  
[PubMed Abstract](#) | [CrossRef Full Text](#) | [Google Scholar](#)
- Liu, W., Kim, M. Y., Van, K., Lee, Y., Li, H., Liu, X., et al. (2011). QTL identification of yield-related traits and their association with flowering and maturity in soybean. *J. Crop Sci. Biotechnol.* 14 (1), 65–70. doi: 10.1007/s12892-010-0115-7  
[CrossRef Full Text](#) | [Google Scholar](#)
- Liu, N., Li, M., Hu, X., Ma, Q., Mu, Y., Tan, Z., et al. (2017). Construction of high-density genetic map and QTL mapping of yield-related and two quality traits in soybean RILs population by RAD-sequencing. *BMC Genomics* 18 (1), 466. doi: 10.1186/s12864-017-3854-8  
[PubMed Abstract](#) | [CrossRef Full Text](#) | [Google Scholar](#)
- Liu, D., Yan, Y., Fujita, Y., Xu, D. (2018). Identification and validation of QTLs for 100-seed weight using chromosome segment substitution lines in soybean. *Breed. Sci.* 68 (4), 442–448.  
doi: 10.1270/jsbbs.17127  
[PubMed Abstract](#) | [CrossRef Full Text](#) | [Google Scholar](#)
- Li, N., Xu, R., Li, Y. (2019). Molecular networks of seed size control in plants. *Annu. Rev. Plant Biol.* 70 (1), 435–463. doi: 10.1146/annurev-arplant-050718-095851  
[PubMed Abstract](#) | [CrossRef Full Text](#) | [Google Scholar](#)
- Li, W., Zheng, D., Van, K., Lee, S. (2008). QTL mapping for major agronomic traits across two years in soybean (*Glycine max* L. merr.). *J. Crop Sci. Biotechnol.* 11 (3), 171–176.  
[Google Scholar](#)

- Mansur, L. M., Lark, K. G., Kross, H., Oliveira, A. (1993). Interval mapping of quantitative trait loci for reproductive, morphological, and seed traits of soybean (*Glycine max* L.). *Theor. Appl. Genet.* 86 (8), 907–913. doi: 10.1007/BF00211040  
[PubMed Abstract](#) | [CrossRef Full Text](#) | [Google Scholar](#)
- Mao, T., Jiang, Z., Han, Y., Teng, W., Zhao, X., Li, W. (2013). Identification of quantitative trait loci underlying seed protein and oil contents of soybean across multi-genetic backgrounds and environments. *Plant Breeding.* 132 (6), 630–641. doi: 10.1111/pbr.12091  
[CrossRef Full Text](#) | [Google Scholar](#)
- Maughan, P. J., Maroof, M. A., Buss, G. R. (1996). Molecular-marker analysis of seed-weight: Genomic locations, gene action, and evidence for orthologous evolution among three legume species. *Theor. Appl. Genet.* 93 (4), 574–579. doi: 10.1007/BF00417950  
[PubMed Abstract](#) | [CrossRef Full Text](#) | [Google Scholar](#)
- Mccouch, S., Cho, Y., Yano, M., Paul, E., Blinstrub, M., Morishima, H., et al. (1997). Report on QTL nomenclature. *Rice Genet. Newsl.* 14, 11–13.  
[Google Scholar](#)
- Mian, M. A., Bailey, M. A., Tamulonis, J. P., Shipe, E. R., Carter, T. E. J., Parrott, W. A., et al. (1996). Molecular markers associated with seed weight in two soybean populations. *Theor. Appl. Genet.* 93 (7), 1011–1016. doi: 10.1007/BF00230118  
[PubMed Abstract](#) | [CrossRef Full Text](#) | [Google Scholar](#)
- Ning, H., Yuan, J., Dong, Q., Li, W., Xue, H., Wang, Y., et al. (2018). Identification of QTLs related to the vertical distribution and seed-set of pod number in soybean [*Glycine max* (L.) merri]. *PLoS One* 13 (4), e195830. doi: 10.1371/journal.pone.0195830  
[CrossRef Full Text](#) | [Google Scholar](#)
- Niu, Y., Xu, Y., Liu, X., Yang, S., Wei, S., Xie, F., et al. (2013). Association mapping for seed size and shape traits in soybean cultivars. *Mol. Breeding.* 31 (4), 785–794. doi: 10.1007/s11032-012-9833-5  
[CrossRef Full Text](#) | [Google Scholar](#)
- Nyquist, W. E., Baker, R. J. (1991). Estimation of heritability and prediction of selection response in plant populations. *Crit. Rev. Plant Sci.* 10 (3), 235–322.  
doi: 10.1080/07352689109382313  
[CrossRef Full Text](#) | [Google Scholar](#)
- Palomeque, L., Li-Jun, L., Li, W., Hedges, B., Cober, E. R., Rajcan, I. (2009). QTL in mega-environments: II. agronomic trait QTL co-localized with seed yield QTL detected in a population derived from a cross of high-yielding adapted × high-yielding exotic soybean lines. *Theor. Appl. Genet.* 119 (3), 429–436. doi: 10.1007/s00122-009-1048-8  
[PubMed Abstract](#) | [CrossRef Full Text](#) | [Google Scholar](#)
- Panthee, D. R., Pantalone, V. R., West, D. R., Saxton, A. M., Sams, C. E. (2005). Quantitative trait loci for seed protein and oil concentration, and seed size in soybean. *Crop Sci.* 45 (5), 2015–2022. doi: 10.2135/cropsci2004.0720  
[CrossRef Full Text](#) | [Google Scholar](#)
- Pathan, S. M., Vuong, T., Clark, K., Lee, J., Shannon, J. G., Roberts, C. A., et al. (2013). Genetic mapping and confirmation of quantitative trait loci for seed protein and oil contents and seed weight in soybean. *Crop Sci.* 53 (3), 765–774. doi: 10.2135/cropsci2012.03.0153  
[CrossRef Full Text](#) | [Google Scholar](#)
- Qi, Z., Wu, Q., Han, X., Sun, Y., Du, X., Liu, C., et al. (2011). Soybean oil content QTL mapping and integrating with meta-analysis method for mining genes. *Euphytica.* 179 (3), 499–514. doi: 10.1007/s10681-011-0386-1  
[CrossRef Full Text](#) | [Google Scholar](#)

- Qi, Z. M., Zhang, X. Y., Qi, H. D., Xin, D. W., Han, X., Jiang, H. W., et al. (2017). Identification and validation of major QTLs and epistatic interactions for seed oil content in soybeans under multiple environments based on a high-density map. *Euphytica*. 213 (8), 162. doi: 10.1007/s10681-017-1952-y  
[CrossRef Full Text](#) | [Google Scholar](#)
- Salas, P., Oyarzo-Llaipen, J. C., Wang, D., Chase, K., Mansur, L. (2006). Genetic mapping of seed shape in three populations of recombinant inbred lines of soybean (*Glycine max* l. merr.). *Theor. Appl. Genet.* 113 (8), 1459–1466. doi: 10.1007/s00122-006-0392-1  
[PubMed Abstract](#) | [CrossRef Full Text](#) | [Google Scholar](#)
- Sax, K. (1923). The association of size differences with seed-coat pattern and pigmentation in PHASEOLUS VULGARIS. *Genetics*. 8 (6), 552–560. doi: 10.1093/genetics/8.6.552  
[PubMed Abstract](#) | [CrossRef Full Text](#) | [Google Scholar](#)
- Sebolt, A. M., Shoemaker, R. C., Diers, B. W. (2000). Analysis of a quantitative trait locus allele from wild soybean that increases seed protein concentration in soybean. *Crop Sci.* 40 (5), 1438–1444. doi: 10.2135/cropsci2000.4051438x  
[CrossRef Full Text](#) | [Google Scholar](#)
- Shomura, A., Izawa, T., Ebana, K., Ebitani, T., Kanegae, H., Konishi, S., et al. (2008). Deletion in a gene associated with grain size increased yields during rice domestication. *Nat. Genet.* 40 (8), 1023–1028. doi: 10.1038/ng.169  
[PubMed Abstract](#) | [CrossRef Full Text](#) | [Google Scholar](#)
- Song, X., Huang, W., Shi, M., Zhu, M., Lin, H. (2007). A QTL for rice grain width and weight encodes a previously unknown RING-type E3 ubiquitin ligase. *Nat. Genet.* 39 (5), 623–630. doi: 10.1038/ng2014  
[PubMed Abstract](#) | [CrossRef Full Text](#) | [Google Scholar](#)
- Sun, L., Li, X., Fu, Y., Zhu, Z., Tan, L., Liu, F., et al. (2013). GS6, a member of the GRAS gene family, negatively regulates grain size in rice. *J. Integr. Plant Biol.* 55 (10), 938–949. doi: 10.1111/jipb.12062  
[PubMed Abstract](#) | [CrossRef Full Text](#) | [Google Scholar](#)
- Sun, Y. Q., Tian, R., Shao, Z. Q., Chen, S. L., Zhang, H., Jin, Y., et al. (2021). Mining of quantitative trait loci and candidate genes for seed size and shape across multiple environments in soybean (*Glycine max*). *Plant Breeding*. 140 (6), 1058–1069. doi: 10.1111/pbr.12968  
[CrossRef Full Text](#) | [Google Scholar](#)
- Tajuddin, T. C. F. T., Watanabe, S., Yamanaka, N., Harada, K. (2003). Analysis of quantitative trait loci for protein and lipid contents in soybean seeds using recombinant inbred lines. *Breed. Sci.* 53 (2), 133–140. doi: 10.1270/jsbbs.53.133  
[CrossRef Full Text](#) | [Google Scholar](#)
- Teng, W., Han, Y., Du, Y., Sun, D., Zhang, Z., Qiu, L., et al. (2009). QTL analyses of seed weight during the development of soybean (*Glycine max* l. merr.). *Heredity*. 102 (4), 372–380. doi: 10.1038/hdy.2008.108  
[PubMed Abstract](#) | [CrossRef Full Text](#) | [Google Scholar](#)
- Teng, W. L., Sui, M. N., Li, W., Wu, D. P., Zhao, X., Li, H. Y., et al. (2018). Identification of quantitative trait loci underlying seed shape in soybean across multiple environments. *J. Agric. Science*. 156 (1), 3–12. doi: 10.1017/S002185961700082X  
[CrossRef Full Text](#) | [Google Scholar](#)
- Wang, H., Jia, J., Cai, Z., Duan, M., Jiang, Z., Xia, Q., et al. (2022). Identification of quantitative trait loci (QTLs) and candidate genes of seed iron and zinc content in soybean [*Glycine max* (L.) merr.]. *BMC Genomics* 23 (1), 146. doi: 10.1186/s12864-022-08313-1  
[PubMed Abstract](#) | [CrossRef Full Text](#) | [Google Scholar](#)

Wang, X., Jiang, G., Green, M., Scott, R. A., Song, Q., Hyten, D. L., et al. (2014). Identification and validation of quantitative trait loci for seed yield, oil and protein contents in two recombinant inbred line populations of soybean. *Mol. Genet. Genomics* 289 (5), 935–949. doi: 10.1007/s00438-014-0865-x

[PubMed Abstract](#) | [CrossRef Full Text](#) | [Google Scholar](#)

Wang, S., Li, S., Liu, Q., Wu, K., Zhang, J., Wang, S., et al. (2015). The OsSPL16-GW7 regulatory module determines grain shape and simultaneously improves rice yield and grain quality. *Nat. Genet.* 47 (8), 949–954. doi: 10.1038/ng.3352

[PubMed Abstract](#) | [CrossRef Full Text](#) | [Google Scholar](#)

Wang, J., Liu, L., Guo, Y., Wang, Y., Zhang, L., Jin, L., et al. (2014). A dominant locus, qBSC-1, controls  $\beta$  subunit content of seed storage protein in soybean (*Glycine max* (L.) merri.). *J. Integr. Agr.* 13 (9), 1854–1864. doi: 10.1016/S2095-3119(13)60579-1

[CrossRef Full Text](#) | [Google Scholar](#)

Warrington, C. V., Abdel-Haleem, H., Hyten, D. L., Cregan, P. B., Orf, J. H., Killam, A. S., et al. (2015). QTL for seed protein and amino acids in the benning  $\times$  danbaekkong soybean population. *Theor. Appl. Genet.* 128 (5), 839–850. doi: 10.1007/s00122-015-2474-4

[PubMed Abstract](#) | [CrossRef Full Text](#) | [Google Scholar](#)

Weng, J., Gu, S., Wan, X., Gao, H., Guo, T., Su, N., et al. (2008). Isolation and initial characterization of GW5, a major QTL associated with rice grain width and weight. *Cell Res.* 18 (12), 1199–1209. doi: 10.1038/cr.2008.307

[PubMed Abstract](#) | [CrossRef Full Text](#) | [Google Scholar](#)

Whiting, R. M., Torabi, S., Lukens, L., Eskandari, M. (2020). Genomic regions associated with important seed quality traits in food-grade soybeans. *BMC Plant Biol.* 20 (1), 485.

doi: 10.1186/s12870-020-02681-0

[PubMed Abstract](#) | [CrossRef Full Text](#) | [Google Scholar](#)

Xian, P., Cai, Z., Jiang, B., Xia, Q., Cheng, Y., Yang, Y., et al. (2022). GmRmd1 encodes a TIR-NBS-BSP protein and confers resistance to powdery mildew in soybean. *Plant Commun.* 3 (6), 100418. doi: 10.1016/j.xplc.2022.100418

[PubMed Abstract](#) | [CrossRef Full Text](#) | [Google Scholar](#)

Xu, Y., Li, H., Li, G., Wang, X., Cheng, L., Zhang, Y. (2011). Mapping quantitative trait loci for seed size traits in soybean (*Glycine max* l. merr.). *Theor. Appl. Genet.* 122 (3), 581–594.

doi: 10.1007/s00122-010-1471-x

[PubMed Abstract](#) | [CrossRef Full Text](#) | [Google Scholar](#)

Yang, K., Moon, J., Jeong, N., Chun, H., Kang, S., Back, K., et al. (2011). Novel major quantitative trait loci regulating the content of isoflavone in soybean seeds. *Genes Genom.* 33 (6), 685–692. doi: 10.1007/s13258-011-0043-z

[CrossRef Full Text](#) | [Google Scholar](#)

Yang, H., Wang, W., He, Q., Xiang, S., Tian, D., Zhao, T., et al. (2017). Chromosome segment detection for seed size and shape traits using an improved population of wild soybean chromosome segment substitution lines. *Physiol. Mol. Biol. Plants.* 23 (4), 877–889.

doi: 10.1007/s12298-017-0468-1

[PubMed Abstract](#) | [CrossRef Full Text](#) | [Google Scholar](#)

Yan, L., Li, Y., Yang, C., Ren, S., Chang, R., Zhang, M., et al. (2014). Identification and validation of an over-dominant QTL controlling soybean seed weight using populations derived from *Glycine max*  $\times$  *Glycine soja*. *Plant Breeding.* 133 (5), 632–637. doi: 10.1111/pbr.12197

[CrossRef Full Text](#) | [Google Scholar](#)

Yao, D., Liu, Z. Z., Zhang, J., Liu, S. Y., Qu, J., Guan, S. Y., et al. (2015). Analysis of quantitative trait loci for main plant traits in soybean. *Genet. Mol. Res.* 14 (2), 6101–6109. doi: 10.4238/2015.June.8.8

[PubMed Abstract](#) | [CrossRef Full Text](#) | [Google Scholar](#)

Yesudas, C. R., Bashir, R., Geisler, M. B., Lightfoot, D. A. (2013). Identification of germplasm with stacked QTL underlying seed traits in an inbred soybean population from cultivars Essex and Forrest. *Mol. Breeding.* 31 (3), 693–703. doi: 10.1007/s11032-012-9827-3

[CrossRef Full Text](#) | [Google Scholar](#)

Yue, Y., Liu, N., Jiang, B., Li, M., Wang, H., Jiang, Z., et al. (2017). A single nucleotide deletion in *J* encoding GmELF3 confers long juvenility and is associated with adaptation of tropic soybean. *Mol. Plant* 10 (4), 656–658. doi: 10.1016/j.molp.2016.12.004

[PubMed Abstract](#) | [CrossRef Full Text](#) | [Google Scholar](#)

Zeng, Z. B. (1994). Precision mapping of quantitative trait loci. *Genet. (Austin)*. 136 (4), 1457–1468. doi: 10.1093/genetics/136.4.1457

[CrossRef Full Text](#) | [Google Scholar](#)

Zhang, J., Song, Q., Cregan, P. B., Jiang, G. (2016). Genome-wide association study, genomic prediction and marker-assisted selection for seed weight in soybean (*Glycine max*). *Theor. Appl. Genet.* 129 (1), 117–130. doi: 10.1007/s00122-015-2614-x

[PubMed Abstract](#) | [CrossRef Full Text](#) | [Google Scholar](#)

Zhang, X., Wang, J., Huang, J., Lan, H., Wang, C., Yin, C., et al. (2012). Rare allele of *OsPPKL1* associated with grain length causes extra-large grain and a significant yield increase in rice. *Proc. Natl. Acad. Sci.* 109 (52), 21534–21539. doi: 10.1073/pnas.1219776110

[CrossRef Full Text](#) | [Google Scholar](#)