



Untargeted liquid chromatography–tandem mass spectrometry metabolomics reveals metabolite inheritance, diversity and capsaicinoid variation across successive generations of *Capsicum chinense* Jacq. (Bhut Jolokia)

Mitali Baruah, Minakshi Bhattacharjee, Robinson C Jose & Rajeev Sarmah*

Programme of Biotechnology, Assam Down Town University, Panikhaiti, Guwahati 781 026, Assam, India

*Correspondence: Rajeev Sarmah, akendrarajeev@gmail.com

Abstract

The *Capsicum chinense* Jacq. pepper is known for its exceptional pungency and rich phytochemical composition. Although metabolite diversity in pepper has been extensively investigated, little is known about metabolite variation across successive generations. This study evaluated the untargeted and targeted metabolite profiles of pepper from two lineages (A3 and A4) across three successive generations through liquid chromatography–tandem mass spectrometry (LC–MS/MS) and high-performance liquid chromatography (HPLC), respectively and analysed their diversity and distributions. A total of 103 metabolites were identified, with 77 metabolites detected in the A3 lineage and 46 in the A4 lineage. Both lineages possessed a conserved core metabolome, while numerous lineage- and generation-specific metabolites demonstrated dynamic metabolic variation during successive propagation. The A3 lineage exhibited greater metabolite richness than the A4 lineage, suggesting enhanced biochemical diversity. Multivariate analyses revealed that most metabolites were shared among generations, whereas a smaller proportion represented unique generation-specific compounds. Clustering analysis clearly separated the samples into two distinct clusters, indicating that genotype was the primary determinant of metabolite composition. Capsaicinoid analysis revealed significant variation in pungency across generations, with the second generation exhibiting the highest Scoville heat unit (SHU) values and the third generation showing a marked reduction. Statistical analysis confirmed significant differences in SHU among generations, particularly between the second and third generations. Furthermore, several capsaicinoid-related metabolites, including homocapsaicin, homodihydrocapsaicin and nordihydrocapsaicin, were detected exclusively in the A3 lineage, highlighting lineage-specific regulation of specialised metabolism. These findings provide new insights into metabolic diversity and inheritance in *C. chinense* and establish a valuable foundation for metabolomics-assisted breeding aimed at improving pungency, nutritional quality and functional phytochemical traits.

Keywords: Bhut Jolokia; capsaicinoids; *Capsicum chinense*; liquid chromatography–tandem mass spectrometry; metabolite inheritance; metabolite profiling; plant breeding; Scoville heat units; successive generations; untargeted metabolomics

Received: 22 July 2026. **Accepted:** 15 September 2026. **Published:** Version 1.0: 03 October 2026; Version 2.0: 13 October 2026.

Cite this article: Mitali B, Minakshi B, Robinson CJ, Rajeev S. Untargeted liquid chromatography–tandem mass spectrometry metabolomics reveals metabolite inheritance, diversity and capsaicinoid variation across successive generations of *Capsicum chinense* Jacq. (Bhut Jolokia). Plant Science Today. 2026; 13(4): 1-21. <https://doi.org/10.14719/pst.16939>

1. Introduction

Capsaicinoids, primarily capsaicin and dihydrocapsaicin, are the principal alkaloids responsible for the characteristic pungency of *Capsicum* fruits. These compounds not only determine fruit quality but also possess considerable pharmaceutical and industrial significance owing to their diverse biological activities (1, 2). A higher accumulation of capsaicinoids is therefore highly desirable for the development of pharmaceutical formulations and functional food products. Studies have demonstrated that capsaicin and related capsaicinoids confer multiple health benefits, including analgesic effects, enhanced energy metabolism, anti-inflammatory

activity, antimicrobial properties, antioxidant potential and free radical scavenging capacity (3). Furthermore, capsaicinoids exhibit a broad spectrum of pharmacological activities, such as anti-obesity, neuroprotective, anticancer, gastroprotective and cardioprotective effects, primarily through modulation of the transient receptor potential vanilloid 1 (TRPV1) signalling pathway, emphasising their therapeutic potential in the prevention and management of pain and various chronic disorders (4). Consequently, increasing capsaicinoid content through conventional breeding, molecular breeding and genetic engineering has become a major focus of *Capsicum* improvement programs (5–7).

In addition to capsaicinoids, *Capsicum* fruits contain a diverse array of primary and secondary metabolites, including phenolic acids, flavonoids, carotenoids, amino acids, organic acids, fatty acids, sugars, vitamins and volatile compounds. These metabolites collectively influence fruit colour, flavour, aroma, nutritional quality and stress tolerance while contributing substantially to the nutraceutical and pharmaceutical value of peppers (8, 9). Comprehensive metabolomic profiling has therefore emerged as an effective strategy for understanding metabolic diversity and identifying biomarkers associated with fruit quality and health-promoting attributes (10).

Capsaicinoid biosynthesis and accumulation are strongly influenced by genotype, environmental conditions, developmental stage, agronomic practices and genotype $\times$ environment interactions (11–13). These factors significantly affect the concentrations of total and individual capsaicinoids, resulting in substantial differences in fruit pungency among cultivars and across growing environments (14). Although several studies have investigated the influence of genotype and environment on capsaicinoid biosynthesis, information regarding metabolite variation across successive generations within a lineage remains scarce. Understanding such generational metabolic variation is important for identifying stable biochemical traits and for developing breeding strategies aimed at improving pungency and fruit quality.

Among cultivated chili peppers, *C. chinense* Jacq., particularly the landraces indigenous to Northeast India, is renowned for its exceptionally high pungency. These landraces, including Bhut Jolokia, are traditionally cultivated in homestead farming systems where seeds are farmer-selected and conserved over successive generations, providing opportunities for gradual genetic and biochemical variation (15). Bhut Jolokia has attracted worldwide attention because of its extraordinarily high capsaicinoid content and its increasing applications in the food, pharmaceutical, cosmetic and defence industries (16).

In the present study, a field experiment was conducted to investigate biochemical variation across successive generations of *C. chinense* (Bhut Jolokia). Untargeted high-resolution liquid chromatography–tandem mass spectrometry (LC–MS/MS) was employed to comprehensively characterise the metabolomic profiles of fruits from different generations. In parallel, capsaicin and dihydrocapsaicin were quantitatively determined using high-performance liquid chromatography (HPLC). The integrated analytical approach was aimed at gaining insights into generational metabolic changes.

2. Materials and Methods

Fruit samples

In continuation with the previous study of morpho-agronomic variations of Bhut Jolokia in successive

generations. The current study targeted insights into metabolomic variations in generations; thus, the experimental design and crop management were the same as earlier research work (17). The plants of Bhut Jolokia were cultivated and managed in the same field through the 2022–25 growing seasons (June to February) and temperature and rainfall conditions were recorded throughout each growing season.

For untargeted LC-MS/MS, 15 matured and ripened fruits (bright red) were collected from plants with a high yield, from both the lineages (Two) in respective generations (Three), accounting for a total sample size of six (plant age $\approx$ 195–210 days). Composite biological samples were prepared with fruits from three high-yielding plants except F1 by drying them at 450 °C for 96 hr. The dried fruits were powdered using a grinder and stored at 40 °C until further use.

Sample preparation for liquid chromatography–tandem mass spectrometry and high-performance liquid chromatography

In a conical flask, 5 g of chili powder was dissolved in 50 mL of methanol and sonicated three times (Rivotek Ultrasonic Cleaner, 50 W, 50 Hz) at 25 °C for 15 min each time. The extracts were then filtered through Whatman No. 1 filter paper and vacuum-evaporated using a rotary evaporator at 40 °C. The resulting dried extracts 1 mg was reconstituted in LC–MS grade methanol to a concentration of 1 ppm (parts per million) and stored at 4 °C until analysis (18) and profiling was carried out using LC-MS/MS facility of the Sophisticated Analytical Instrument Facility (SAIF), IIT Bombay.

Similarly, for quantification of capsaicin and dihydrocapsaicin, 1 g of chili powder was dissolved in 10 mL of acetonitrile. One milligram (1 mg) of the dried sample was reconstituted in 1 mL of HPLC-grade methanol, resulting in a final concentration of 1 mg/mL. The quantification was performed using the HPLC facility of the Biotech Park, Guwahati.

Metabolite profiling

Metabolite Profiling was performed using an Agilent 1290 Infinity UHPLC system coupled to an Agilent 6550 Quadrupole Time of Flight (QTOF) mass spectrometer (Agilent Technologies, USA) equipped with an electrospray ionisation (ESI) source. Chromatographic separations were achieved on a Zorbax Eclipse Plus C18 column (2.1 $\times$ 100 mm, 1.8 μ m) maintained at 35 °C. The mobile phases consisted of water with 0.1 % formic acid (solvent A) and acetonitrile with 0.1 % formic acid (solvent B). A linear gradient was applied as follows: 5 % B (0–2 min), 5–95 % B (2–20 min), 95 % B (20–25 min), followed by re-equilibration to initial conditions. The flow rate was 0.3 mL/min and the injection volume was 3 μ L.

The QTOF was operated in both positive and negative ESI modes under the following conditions: capillary voltage 3.5 kV, drying gas flow 10 L/min at 325 °C, nebuliser pressure 35 psi and fragmentor voltage 120 V. Full-scan data were acquired over an m/z range



of 50–1200 with an acquisition rate of 2 spectra/s. Reference ions at m/z 121.0509 and 922.0098 were continuously infused to ensure mass accuracy.

The LC-MS/MS data was processed using Agilent MassHunter Qualitative Analysis software. Molecular features were extracted based on retention time, accurate mass and isotope pattern. Metabolite features detected through untargeted LC-MS/MS analysis were putatively annotated by comparing the acquired LC-MS/MS data with the METLIN database (19). The plant-based compounds were confirmed through a literature search and presented.

Quantification of capsaicin and dihydrocapsaicin

The reagents used were of HPLC grade. Standard capsaicin (HiMedia, 50 % capsaicin and 40 % dihydrocapsaicin) was used for instrument calibration. Aqueous solutions were prepared with de-ionised water obtained from the Ultrapure water system of Sartorius Stedim Biotech, Germany, throughout the experiments. All the standards and samples were kept cool at 4 °C before use.

The experiment was carried out on HPLC system (Agilent, 1260 Infinity Series) equipped with a. HPLC pump (1260 Quat Pump) b. C18 RP Column (4.6 × 150 mm) and packed with particles c. Dual λ absorbance detector (1260 MWDVL) with detection at 280 nm. The system consists of a Rheodyne Injector with a sample loop of 20 μ L. OpenLab CDS (Chem station edition) software was used for data analysis.

The extract was dissolved in methanol to make a final volume of 5 mL and the aliquot of this sample was filtered through 0.45 μ m PTFE membrane using a 5 mL disposable syringe and stored in 1.8 mL storage vial. A 20 μ L aliquot was used for each HPLC injection. The HPLC operating conditions to determine total heat units (designated as “short” run) included ambient temperature, a flow rate of 1 mL/min with a run duration of 10 min. The mobile phase was isocratic, with 70 % solvent B (100 % methanol) and 30 % solvent A (10 % methanol).

Standards of capsaicin (8-methyl-n-vanillyl-6-nonenamide) and dihydrocapsaicin (8-methyl-n-vanillyl-nonanamide) were used for retention-time (RT) verification along with instrument calibration. Standard solutions of 1000, 500, 250, 100 and 50 PPM were prepared in 100 % methanol by dilution of a 2000 PPM stock solution.

Estimation of Scoville heat units

Scoville heat units are estimated in parts per million of heat (ppmH) based on the sample's dry weight using the ASTA (1985) method using the formula below.

$$\text{PpmH} =$$

$$\frac{\text{Peak area of Capsaicin} + [(0.82) \times (\text{Peak area of Dihydrocapsaicin})] \times (\text{ppm Standard}) \times (\text{mL acetonitrile})}{$$

$$(\text{Total Capsaicin Peak area of Standard}) \times (\text{g Sample})$$

$$\text{SHU} = \text{ppmH} \times 15$$

Where *ppmH* is the parts per million of heat, calculated on a dry-weight basis, 0.82 is relative heat factor for dihydrocapsaicin, SHU is Scoville Heat Unit and 15 is the mathematical conversion factor. Conversion to micrograms (μ g) total capsaicinoid per gram of sample was done by using the AOAC Official Method 995.03 (AOAC, 1995). 1 Microgram (μ g) total capsaicinoid/g ~ ca 15 SHU (20).

The F_1 , F_2 and F_3 samples were analysed in separate analytical runs using their respective calibration standards and corresponding calibration parameters.

Statistical analysis

The metabolomic and capsaicinoid data were analysed using a Python-based workflow. Metabolite presence–absence patterns were evaluated using Jaccard similarity (21), while normalised metabolite abundances were used for hierarchical clustering, principal component analysis (PCA) (22, 23) and fold-change analysis. The statistical analysis of SHU values across generations was performed using Dunn's multiple-comparison test with Bonferroni correction. Environmental covariates were considered in the adjusted log-abundance analysis (24).

3. Results

Metabolite profile

In total, 103 compounds were identified across the A3 and A4 lineages, filtered based on plant origin through literature. The presence–absence matrix (Supplementary Table 1, Supplementary Fig 1 and 2) illustrates distribution of these metabolites across successive generations of the A3 lineage (A303, A303-06 and A303-06-18) and the A4 lineage (A425, A425-10 and A425-10-15). The mass accuracy and database support for the metabolite annotations are shown in Fig. 1.

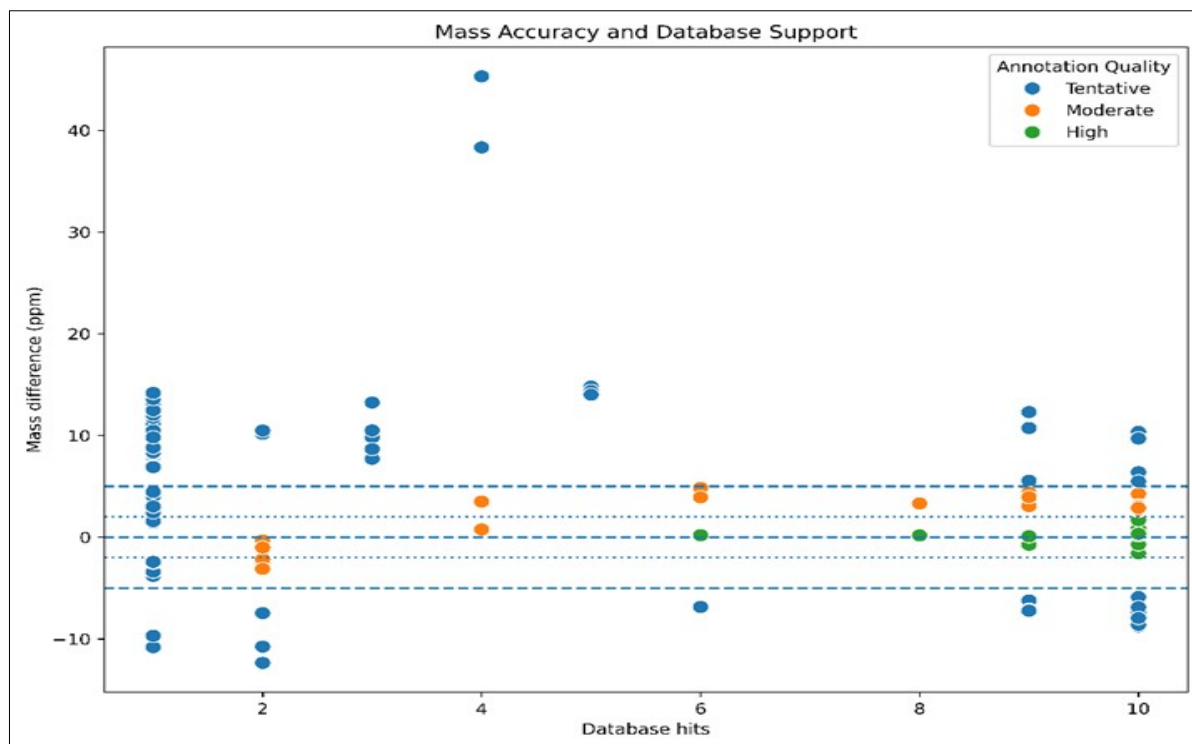


Fig. 1. Mass accuracy and database support of the putatively annotated metabolites, categorised according to annotation quality (tentative, moderate and high).

The distribution of the most abundant metabolites is shown in Supplementary Fig. 3, with phenylacetic acid and capsi-amide exhibiting the highest maximum abundances.

The metabolites presence-absence patterns, showed occurrence of highly distinct biosynthesis inter-lineage, however intra-lineage difference was less to moderate (Fig. 2). The generations of A3 lineage, showed moderate similarity, with coefficients ranging between 0.46 and 0.56, where A303 and A303-06-18 (0.56) shared more common compounds, than A303-06 and A303-06-18 (0.48) and A303 and A303-06 (0.46). In contrast, A4 exhibited higher within-lineage similarity, ranging between 0.53 and 0.77 and A425-10 and A425-10-15 (0.77) had more common compounds. While comparing both the lineages the similarity was considerably low (0.21–0.28), indicating a distinct metabolite profile between them.

The normalised abundance values (Z- Score) based clustering of the metabolite profiles demonstrated existence of lineage-specific pattern (Fig. 3). The two clusters of the generations clearly showed the distinctive pattern of metabolite synthesis, where A425, A425-10 and A425-10-15 clustered together and A303, A303-06 and A303-06-18 formed a separate major cluster. Several compounds contributed to the observed differences in metabolite distribution. Compounds such as 2S-hydroxyoctadecanoic acid, 16-hydroxyhexadecanoic

acid, all-trans-heptaprenyl diphosphate, petroselinic acid, capsoside A, 9-hydroxytridecyl docosanoate, 8S-HODE, octyl octanoate, rhodoxanthin and linayl caprylate showed distinct abundance patterns across the A3 and A4 lineages. Whereas 10'-apo-beta-carotenal, 3-hydroxy-beta-e-caroten-3-one, 2-butyl-4-ethyl-5-methyloxazole, 7,8-dehydro-beta micropteroxanthin and 3-keto stearic acid displayed lineage- and generation-dependent variation.

A group of capsaicinoid-related compounds, including capsi-amide, homocapsaicin, nordihydrocapsaicin, palmitic amide and homodihydrocapsaicin, showed a comparatively similar pattern among the A4 lineage, particularly A425-10 and A425-10-15, consistent with their high Jaccard similarity. In contrast, compounds such as 10'-apo-beta-carotenal, 3-hydroxy-beta-e-caroten-3-one and 2-butyl-4-ethyl-5-methyloxazole contributed to the differentiation observed among lineages. This clustering pattern indicates that metabolite composition was more conserved across successive generations of the A4 lineage than in the A3 lineage, while the two lineages exhibited pronounced differences in their metabolite profiles.

The PCA plot (Fig. 4) further supports the Jaccard similarity and hierarchical clustering analysis, representing a clearer separation between the A3 and A4 lineages based on their metabolite profiles.

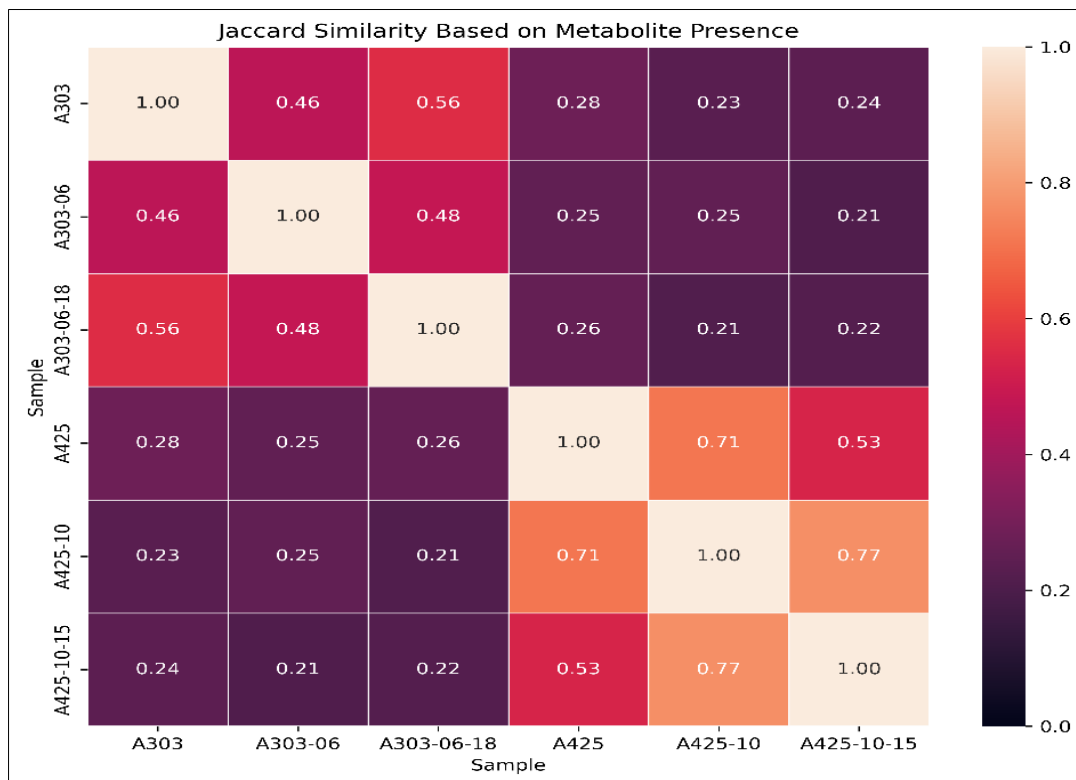


Fig. 2. Jaccard similarity heatmap based on metabolite presence / absence, showing higher intra-lineage metabolite similarity than inter lineages.

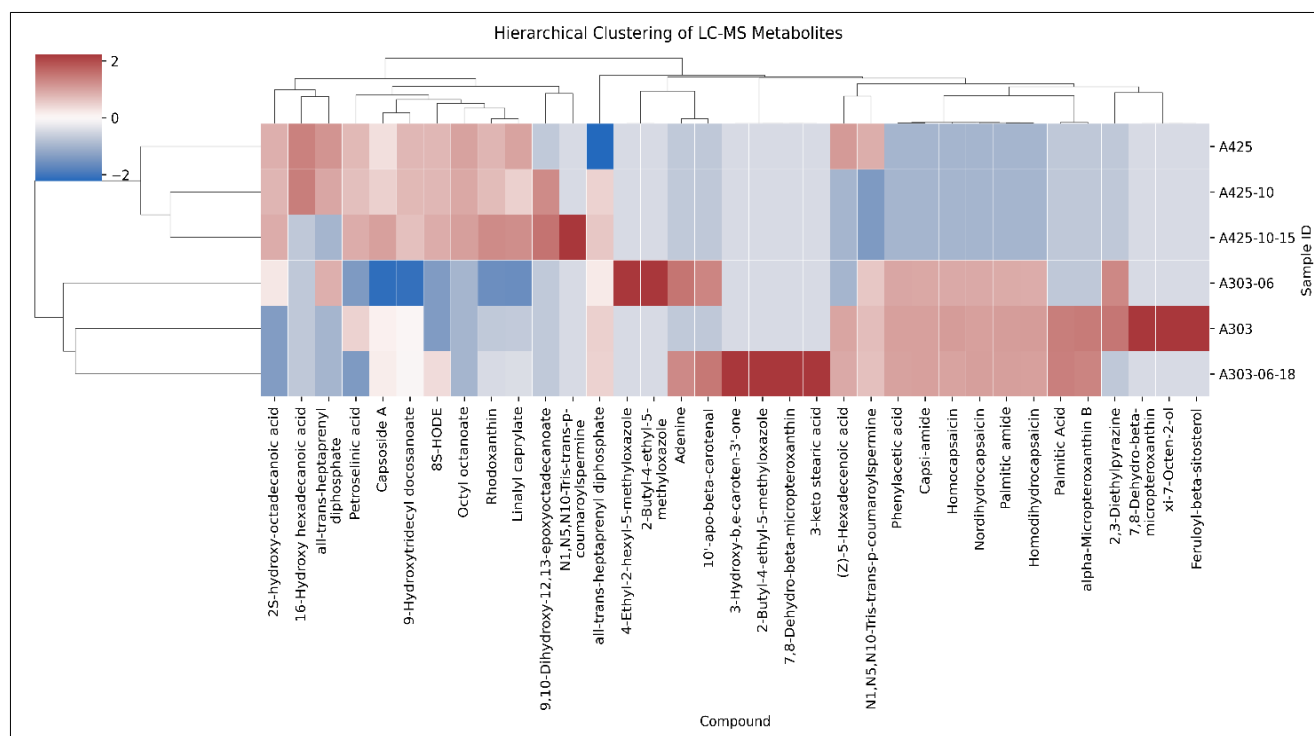


Fig. 3. Hierarchical clustering of metabolites showing distinct separation between the A3 and A4 lineages, with samples from the same lineage displaying greater similarity in their metabolite profiles and illustrated lineage-specific metabolic variation across successive generations.

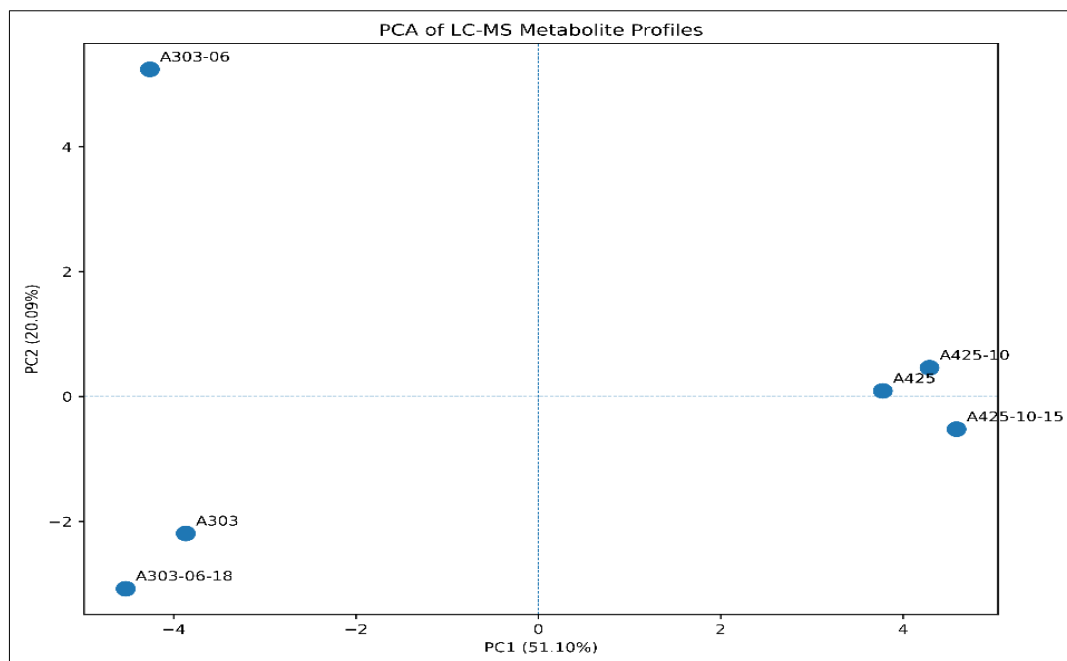


Fig. 4. Principal component analysis (PCA) of liquid chromatography–tandem mass spectrometry (LC–MS/MS) metabolite profiles revealing clear separation between the A3 and A4 lineages. PC1 and PC2 explained 51.10 % and 20.09 % of the total variation, respectively, accounting for 71.19 % cumulative variance. A3 samples clustered on the negative PC1 axis, whereas A4 samples clustered on the positive PC1 axis, indicating distinct metabolite profiles between the two lineages.

Further, PC1 explained 51.10 % and PC2 explained 20.09 % of the total variation, together accounting for 71.19 % of the variance. The generations of A3 (A303, A303-06 and A303-06-18) were positioned on the negative side of PC1, whereas the A4 lineage (A425, A425-10 and A425-10-15) clustered on the positive side. Within A3, A303-06 was distinctly separated along PC2, while A303 and A303-06-18 were closely positioned. The A4 generations clustered relatively closely, indicating greater similarity in their metabolite profiles. It demonstrates a distinct metabolomic differentiation between the A3 and A4 lineages, with comparatively consistent profiles across successive A4 lineages.

The fold-change analysis revealed a substantial generation-dependent variation in metabolite abundance in both lineages. In the A3 lineage (Fig. 5), A303-06 and A303-06-18 exhibited pronounced changes relative to A303, with strong increases in compounds such as adenine, 7,8-dehydro- β -microperoxanthin, 10'-apo- β -carotenal, 2-butyl-4-ethyl-5-methyloxazole, 8S-HODE, 3-keto stearic acid and 3-hydroxy- β , ϵ -caroten-3'-one, while feruloyl- β -sitosterol, xi-7-octen-2-ol, (Z)-5-hexadecenoic acid, petroselinic acid, 9-hydroxytridecyl docosanoate and palmitic acid showed marked decreases.

In the A4 lineage (Fig. 6), pronounced increases were observed for 9,10-dihydroxy-12,13-epoxyoctadecanoate and N1,N5,N10-tris-trans-p-coumaroylspermidine, whereas N1,N5,N10-tris-trans-p-coumaroylspermidine, 16-hydroxyhexadecanoic acid and (Z)-5-hexadecenoic acid showed strong decreases. These results indicated substantial metabolic reprogramming across successive generations, with distinct fold-change patterns between the A3 and A4 lineages.

Above all it was also observed that after controlling for environmental covariates (rainfall, temperature and humidity), the adjusted log-abundance profiles showed clear differences in observation distributions across generation and lineages. Within the A3 lineage, A303 and A303-06-18 had relatively broad distributions, indicating greater within-sample variability, whereas A303-06 had a comparatively narrower distribution. Across the three A3 samples, the median adjusted log-abundance was approximately 11.1–11.4. Whereas, within the A4 lineage, distributions were relatively narrower, with median values increasing from about 11.6 in A425 to 11.9 in A425-10 and 12.1 in A425-10-15. Consequently, A425-10-15 had the highest median adjusted log-abundance among all. This analysis suggest that variation related to lineage and generation persisted even after accounting for environmental covariates, with A3 showing greater variability and A4 displaying a steady upward trend in median abundance across generations (Fig. 7)

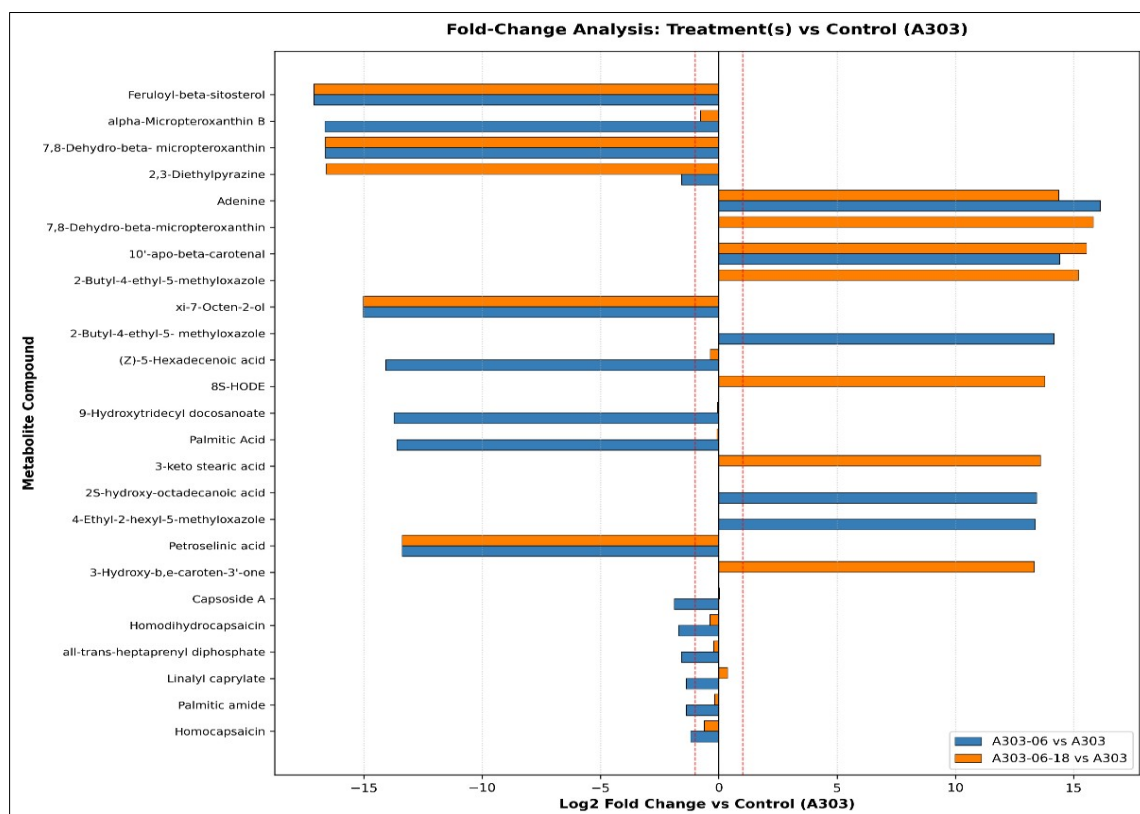


Fig. 5. Fold-change analysis of A3 lineage revealing substantial metabolite variation across successive A3 generations. A303-06 and A303-06-18 showed strong increases in metabolites such as adenine, 7,8-dehydro- β -microperoxanthin, 10'-apo- β -carotenal and 2-butyl-4-ethyl-5-methyloxazole, while several fatty acids and related compounds showed marked decrease.

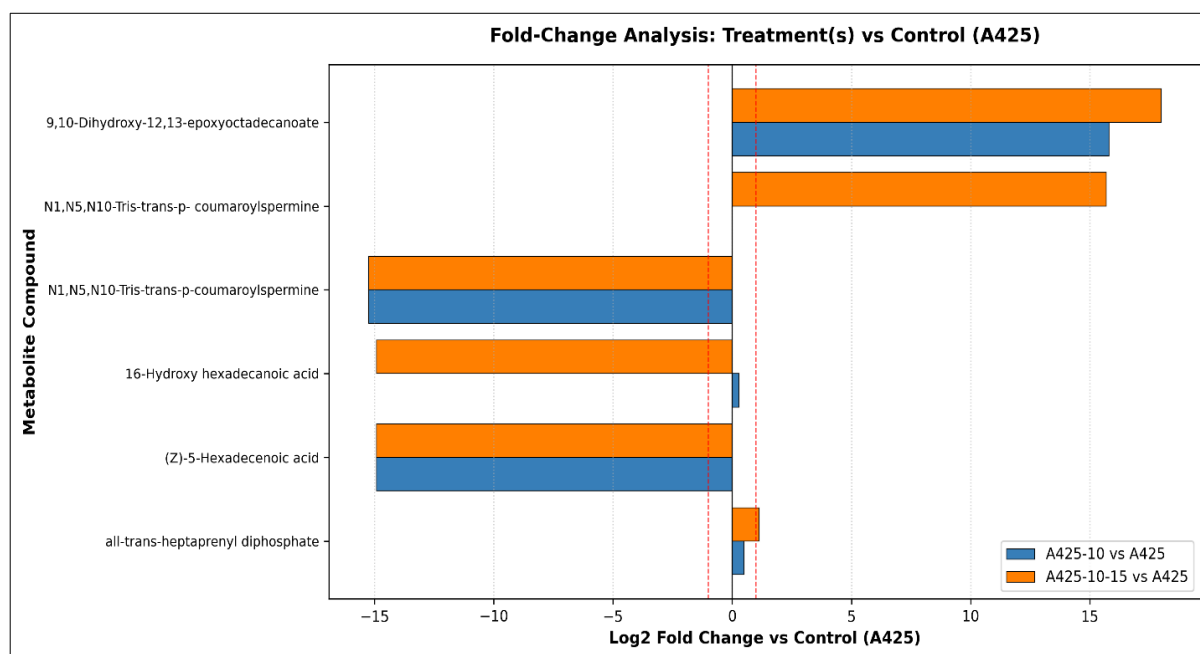


Fig. 6. Fold-change analysis of A4 lineage showing marked metabolic changes across successive generations. 9,10-Dihydroxy-12,13-epoxyoctadecanoate showed the strongest increase, while N1,N5,N10-tris-trans-p-coumaroylspermidine and (Z)-5-hexadecenoic acid showed pronounced decrease.

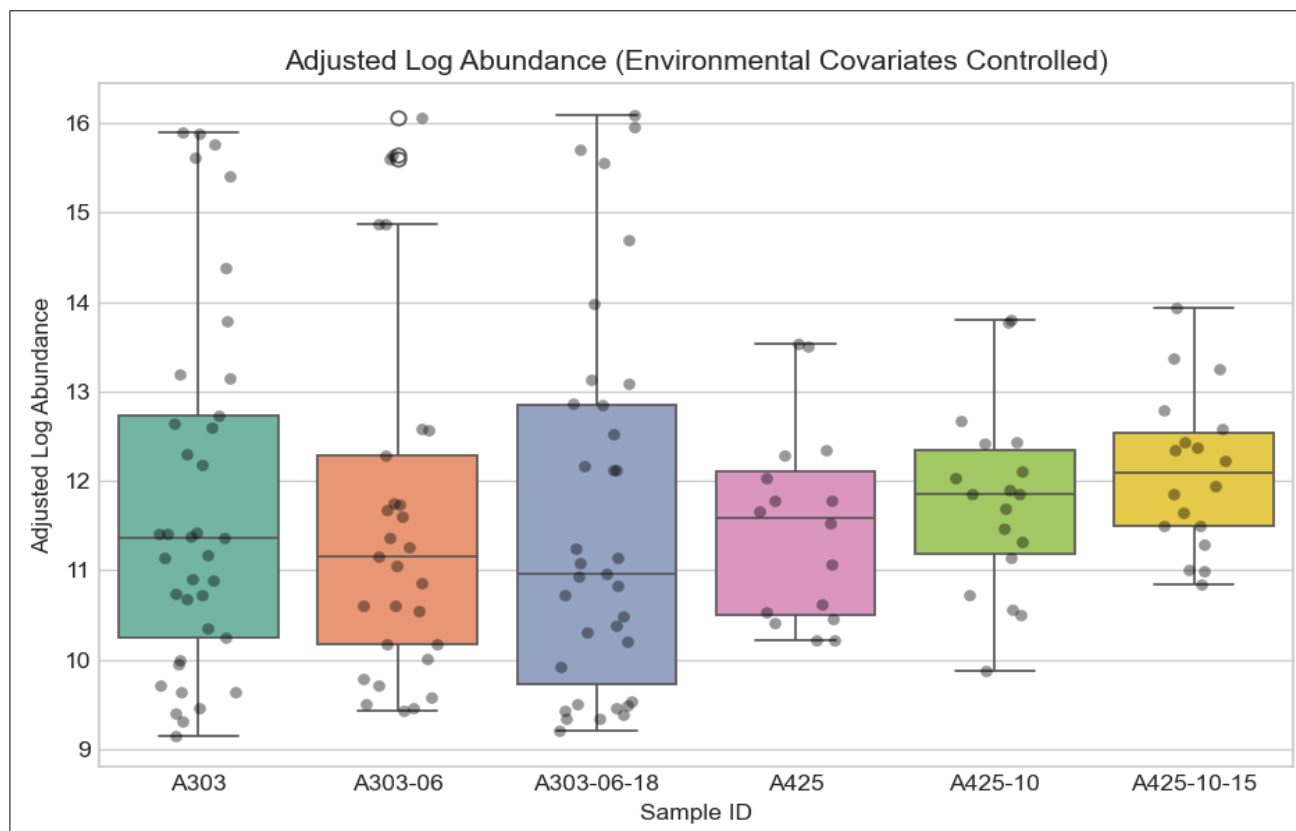


Fig. 7. Adjusted log abundance across sample IDs after controlling for environmental covariates.

A303, A303-06 and A303-06-18 represent the A3 lineage generations F₁, F₂ and F₃, respectively, while A425, A425-10 and A425-10-15 represent the A4 lineage generations F₁, F₂ and F₃, respectively. Boxplots represent the median, interquartile range and whiskers, with individual points showing metabolite observations. Log-abundance values were adjusted for the measured environmental covariates.

Capsaicinoids variations within and across generations

The distribution of SHU among the 3 generations (F₁, F₂ and F₃) differed markedly among the 3 generations. In the A3 lineage, F₂ exhibited the highest pungency, followed by F₁, whereas F₃ had the lowest SHU values, indicating a substantial reduction in capsaicinoid content in the third generation. F₁ showed SHU values ranging from 297000 to 454000, with a median of 377000. A moderate interquartile range (IQR) indicated moderate variability among the three replicates, with no outliers. F₂ showed the highest SHU values, with a median of 690000, concentrated between 650000 and 730000 and one observation of 197000 SHU identified as a low outlier. Overall, F₂ consistently exhibited substantially greater pungency than F₁ and F₃. F₃ had the lowest SHU values, ranging from 75000 to 145000 SHU, with a median of 95000 SHU.

(Fig. 8, Supplementary Fig. 4 and 5). The box in the figure is relatively narrow, indicating comparatively low variability among replicates. These results suggest a pronounced decline in capsaicinoid concentration in the third generation. The medians show a clear pattern: F₂ > F₁ > F₃. A p -value < 0.05 indicates that differences in SHU among the second and third generations are statistically significant. The Dunn's post hoc test (Table 1 and 2, Supplementary Table 2) with Bonferroni-adjusted p -values reflected that SHU differed significantly in F₂ vs F₃.

In the A4 lineage, SHU values varied among the three generations. F₂ exhibited the highest pungency, with SHU values substantially exceeding those of F₁ and F₃. F₁ displayed intermediate SHU values, whereas F₃ showed the lowest median SHU (Fig. 9), indicated a marked reduction in capsaicinoid content. Two high-value outliers were observed in F₃, suggesting considerable biological variability within this generation. Dunn's multiple comparison test with Bonferroni correction demonstrated a highly significant difference between F₂ and F₃ (adjusted $p = 7.46 \times 10^{-6}$), whereas differences between F₁ and F₂ ($p = 1.000$) and between F₁ and F₃ ($p = 0.063$) were not statistically significant (Table 2, Supplementary Table 2).

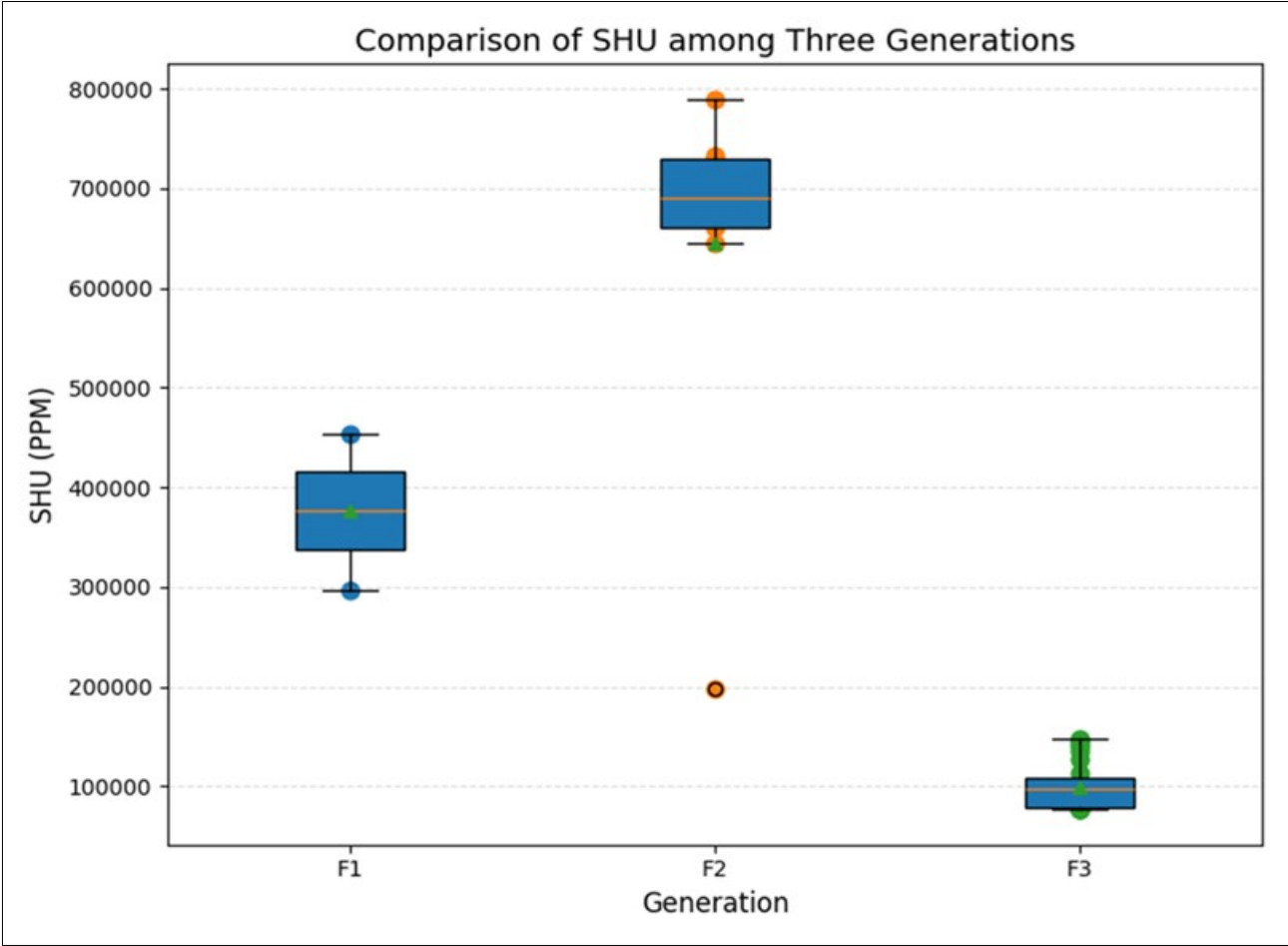


Fig. 8. Boxplot comparing Scoville heat units (SHU) across three successive selfed generations (F1, F2 and F3) of *Capsicum chinense* (Bhut Jolokia).

Table 1. Pairwise comparison of Scoville heat units (SHU) among successive selfed generations of *Capsicum chinense* (Bhut Jolokia) A3 using Dunn's post hoc test

Comparison	Adjusted <i>p</i> -value	Interpretation
F1 vs F2	1.0000	Not significant
F1 vs F3	0.06336	Not significant (<i>p</i> > 0.05)
F2 vs F3	0.00000746	Highly significant (<i>p</i> < 0.001)

Adjusted *p*-values indicate the significance of differences between generations.

Table 2. Pairwise comparison of Scoville heat units (SHU) among successive selfed generations of *Capsicum chinense* (Bhut Jolokia) A4 using Dunn's post hoc test

Comparison	Adjusted <i>p</i> -value	Interpretation
F1 vs F2	1.000	Not significant
F1 vs F3	0.0634	Not significant
F2 vs F3	7.46 × 10 ⁻⁶	Highly significant

Adjusted *p*-values indicate the significance of differences between generations.



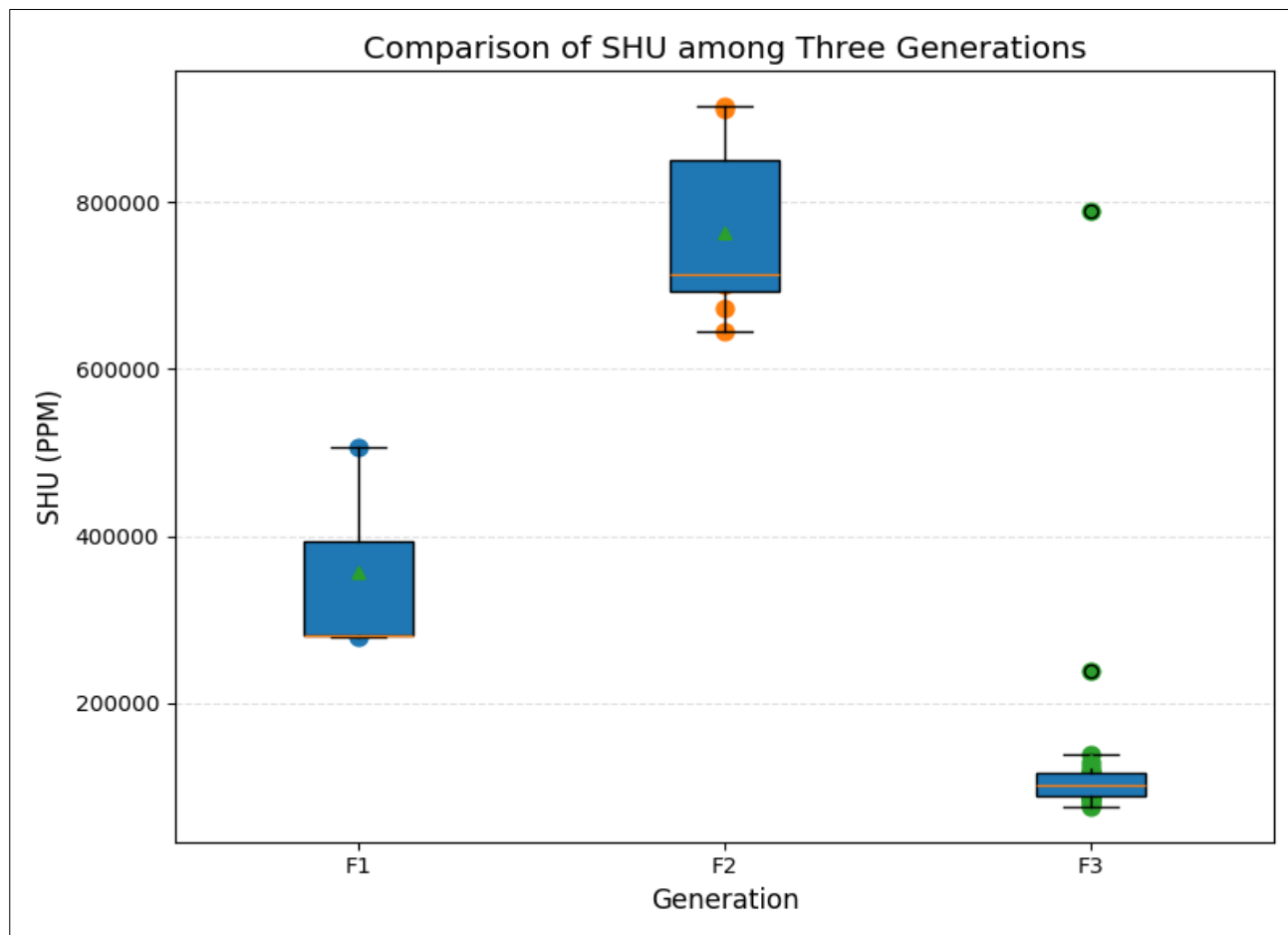


Fig. 9. Boxplot comparing Scoville heat units (SHU) across three successive selfed generations (F1, F2 and F3) of *Capsicum chinense* (Bhut Jolokia).

4. Discussion

Untargeted LC–MS/MS metabolomics revealed substantial differences in metabolite composition between the A3 and A4 lineages of *C. chinense*, while also demonstrating the conservation of a core metabolome across successive generations. The identification of 103 metabolites, with 77 in the A3 and 46 in the A4 lineages, indicates considerable metabolic diversity within lineages and across genotypes. Similar observations have been reported in recent metabolomic studies of *Capsicum*, where genotype has been identified as the principal determinant of metabolite composition, exceeding the influence of environmental variation or developmental stage (25).

The presence–absence matrix demonstrated that several metabolites were consistently detected across all generations and lineages, representing a conserved metabolic core. Such conserved metabolites likely participate in primary metabolic pathways essential for fruit development, membrane stability, carotenoid biosynthesis and cellular homeostasis. Maintenance of a stable core metabolome across generations has also been reported in tomato and pepper metabolomic studies,

where housekeeping metabolites remain highly conserved despite substantial variation in specialised metabolites (26). The persistence of these metabolites suggests that fundamental physiological functions remain genetically stable during successive propagation.

In contrast, numerous lineage-specific metabolites clearly distinguished the populations of A3 and A4. The A3 lineages contained several metabolites associated with alkaloid, amide, sterol and phenylpropanoid metabolism, including homocapsaicin, homodihydrocapsaicin, nordihydrocapsaicin, capsiamide and feruloyldihydro- β -sitosterol, whereas the A4 lineage possessed a different set of specialised metabolites. Such metabolic divergence is expected because secondary metabolism is strongly controlled by genotype, transcriptional regulation and enzyme activity. Previous genomic studies have demonstrated that polymorphisms in genes regulating capsaicinoid biosynthesis, phenylpropanoid metabolism and fatty acid metabolism produce distinct metabolic fingerprints among *Capsicum* accessions (27). Therefore, the metabolite differences observed between the two lineages most likely reflect their distinct genotype rather than random metabolic variation.

The gradual gain and loss of metabolites among successive generations further indicate that metabolic inheritance is dynamic rather than static. Such variation may result from segregation of alleles controlling specialised metabolism, epigenetic regulation, developmental plasticity, or environmental modulation of biosynthetic pathways. Secondary metabolites are particularly sensitive to regulatory changes because their biosynthesis depends on coordinated expression of multiple enzymes and transcription factors. Recent studies have demonstrated that MYB, bHLH and ERF transcription factors modulate numerous secondary metabolic pathways in *Capsicum*, resulting in generation-dependent variation in metabolite accumulation (28). Consequently, the observed metabolite gains and losses likely reflect alterations in pathway regulation rather than complete activation or loss of metabolic capacity.

Further analysis demonstrated that although each sample possessed unique metabolites, most compounds were shared among closely related generations. The A3 lineage exhibited considerably greater metabolite richness than the A4 lineage, suggesting either broader biosynthetic capacity or greater metabolic plasticity. Higher metabolite richness is generally associated with increased biochemical diversity and enhanced adaptive potential because plants possessing more specialised metabolites can respond to diverse environmental and physiological stresses (29). The relatively smaller number of metabolites detected in the A4 lineage may therefore indicate a more conserved metabolic architecture.

The metabolite-class distribution across the A3 and A4 lineages showed considerable chemical diversity across successive generations. In the A3 lineage, metabolites spanned several classes, with other/unclassified metabolites most prominent, followed by fatty acids/oxylipins and alkaloids. The A4 lineage also exhibited a diverse metabolic profile, with lipid-related, terpenoid, alkaloid and other secondary metabolite classes represented. Overall, the presence of multiple metabolite classes in both lineages suggests that successive generations were associated with changes in several metabolic pathways rather than a shift in a single metabolite group. The predominance of unclassified compounds further indicates that a substantial proportion of the detected metabolic diversity remains to be characterised. Similar transient metabolomic shifts have been reported across successive breeding generations in several crop species and are believed to arise from recombination and differential regulation of secondary metabolic pathways before stabilisation in later generations (30).

Hierarchical clustering and principal compound analysis provided additional evidence that metabolite composition primarily reflects genetic lineage. The clear separation of the A3 and A4 lineages, together with the relatively large linkage distance between clusters, demonstrated that overall metabolic composition differs substantially between these populations. Within each

lineage, successive generations clustered closely together, indicating that inheritance preserves much of the parental metabolome despite generation-specific modifications. Similar clustering patterns have been widely observed in untargeted metabolomic studies of pepper, tomato and other Solanaceae species, where genotype consistently explains the largest proportion of metabolic variance (31). The close association of A303 with A303-06-18 further suggests that some transient metabolic alterations present in the intermediate generation may not persist in later generations.

Capsaicinoid analysis revealed substantial variation in pungency among successive generations. Both lineages exhibited maximum SHU values in the second generation followed by a marked decline in the third generation. Because capsaicinoids are synthesised through the convergence of the phenylpropanoid and branched-chain fatty acid pathways, their accumulation depends upon coordinated expression of genes including Pun1, pAMT, Kas, BCAT and the transcription factor CaMYB31. Recent molecular studies have confirmed that variation in expression of these genes directly influences capsaicinoid accumulation and pungency among *Capsicum* cultivars (32). The elevated SHU observed in the second generation may therefore reflect enhanced expression of capsaicinoid biosynthetic genes, whereas the reduced pungency in the third generation may result from segregation of favorable alleles or reduced transcriptional activity.

The statistical analyses further support these biological interpretations. The significant differences observed among generations in the A3 lineage indicated that capsaicinoid accumulation changes substantially during successive propagation. In the A4 lineage, however, only the comparison between F₂ and F₃ was statistically significant following Bonferroni correction. This finding suggests that although the second generation consistently accumulated higher capsaicinoid concentrations, the parental generation maintained intermediate pungency that did not differ significantly from either extreme. Such results are consistent with quantitative inheritance of pungency, where multiple loci contribute additively to capsaicinoid biosynthesis.

An important observation from the metabolomic analysis is the occurrence of homocapsaicin, homodihydrocapsaicin and nordihydrocapsaicin exclusively within the A3 lineage. These compounds are minor capsaicinoids that contribute to overall pungency and possess antioxidant and pharmacological activities comparable to capsaicin. Their absence from the A4 lineage may partially explained differences in total SHU values and indicated lineage-specific regulation of capsaicinoid biosynthesis. Previous metabolomic investigations have similarly shown that minor capsaicinoids vary considerably among *Capsicum* genotypes despite relatively conserved capsaicin biosynthesis (33).

Overall, the present study demonstrated successive generations of *C. chinense* retains a conserved core metabolome while simultaneously undergoing dynamic alterations in specialised metabolism. The strong lineage-specific clustering, differential metabolite richness, generation-dependent accumulation of unique metabolites and significant variation in capsaicinoid content collectively indicate that inheritance influences both qualitative and quantitative aspects of metabolism. The limitation of the present study is that untargeted LC–MS/MS metabolomic analysis was based on a single composite biological sample per generation. Therefore, the observed generation-specific presence or absence of metabolites should be interpreted as descriptive patterns in the analysed samples rather than as statistically replicated evidence of metabolite gains or losses across generations. Future studies using multiple independent biological replicates per generation are required to confirm the reproducibility and biological significance of these generation-specific metabolomic differences.

5. Conclusion

This study provides the first comprehensive assessment of metabolite inheritance across successive generations of Bhut Jolokia using untargeted metabolomics integrated with capsaicinoid analysis. The findings improve the understanding of metabolic stability, inheritance and diversification in *Capsicum chinense* and identify metabolomic signatures that may serve as valuable biomarkers for selecting superior breeding lines with enhanced pungency, nutritional quality and functional phytochemical composition. Future studies integrating metabolomics with transcriptomic and genomic analyses will further elucidate the molecular mechanisms governing metabolite inheritance and facilitate marker-assisted breeding for the genetic improvement of Bhut Jolokia.

Acknowledgements

The authors are grateful to Assam Down Town University, Biotech Park, Guwahati, and IIT Mumbai (SAIF/CRNTS IIT Bombay) providing access to instruments for conducting the various analyses.

Authors' contributions

MB¹ carried out the field experiments and data collection. RCJ and RS carried out the statistical analysis and contributed to manuscript revision. MB² conceived of the study and participated in its design and coordination. All authors have read and approved the final manuscript [MB¹ stands for Mitali Baruah and MB² stands for Minakshi Bhattacharjee].

Compliance with ethical standards

Conflict of interest

Authors do not have any conflict of interests to declare.

Ethical issues

None

References

- Wang D, Bosland PW. The genes of *Capsicum*. HortScience. 2006;41(5):1169–87. <https://doi.org/10.21273/HORTSCI.41.5.1169>
- Barbero GF, Liazid A, Azaroual L, Palma M, Barroso CG. Capsaicinoid contents in peppers and pepper-related spicy foods. Int J Food Prop. 2016;19(3):485–93. <https://doi.org/10.1080/10942912.2014.968468>
- Reyes-Escogido MD, Gonzalez-Mondragon EG, Vazquez-Tzompantzi E. Chemical and pharmacological aspects of capsaicin. Molecules. 2011;16(2):1253–70. <https://doi.org/10.3390/molecules16021253>
- Basith S, Cui M, Hong S, Choi S. Harnessing the therapeutic potential of capsaicin and its analogues in pain and other diseases. Molecules. 2016;21(8):966. <https://doi.org/10.3390/molecules21080966>
- Stewart C Jr, Mazourek M, Stellari GM, O'Connell M, Jahn M. Genetic control of pungency in *C. chinense* via the Pun1 locus. J Exp Bot. 2007;58(5):979–91. <https://doi.org/10.1093/jxb/erl243>
- Arce-Rodríguez ML, Ochoa-Alejo N. Biochemistry and molecular biology of capsaicinoid biosynthesis: recent advances and perspectives. Plant Cell Rep. 2019;38(9):1017–30. <https://doi.org/10.1007/s00299-019-02406-0>
- Aza-González C, Núñez-Palenius HG, Ochoa-Alejo N. Molecular biology of capsaicinoid biosynthesis in chili pepper (*Capsicum* spp.). Plant Cell Rep. 2011;30(5):695–706. <https://doi.org/10.1007/s00299-010-0968-8>
- Deepa N, Kaur C, George B, Singh B, Kapoor HC. Antioxidant constituents in some sweet pepper (*Capsicum annuum* L.) genotypes during maturity. LWT Food Sci Technol. 2007;40(1):121–9. <https://doi.org/10.1016/j.lwt.2005.09.016>
- Howard LR, Talcott ST, Brenes CH, Villalon B. Changes in phytochemical and antioxidant activity of selected pepper cultivars (*Capsicum* species) as influenced by maturity. J Agric Food Chem. 2000;48(5):1713–20. <https://doi.org/10.1021/jf990916t>
- Hall RD. Plant metabolomics in a nutshell: potential and future challenges. Annu Plant Rev. 2011;43:1–24. <https://doi.org/10.1002/9781444339956.ch1>
- Bosland PW, Votava EJ, Votava EM. Peppers: vegetable and spice capsicums. 2nd ed. Wallingford: CABI; 2012. <https://doi.org/10.1079/9781845938253.0000>
- Zewdie Y, Bosland PW. Evaluation of genotype, environment, and genotype-by-environment interaction for capsaicinoids in *Capsicum annuum* L. Euphytica. 2000;111(3):185–90. <https://doi.org/10.1023/A:1003837314929>
- Gurung T, Techawongstien S, Suriham B, Techawongstien S. Impact of environments on the accumulation of capsaicinoids



- in *Capsicum* spp. HortScience. 2011;46(12):1576–81. <https://doi.org/10.21273/HORTSCI.46.12.1576>
14. Wahyuni Y, Ballester AR, Sudarmonowati E, Bino RJ, Bovy AG. Metabolite biodiversity in pepper (*Capsicum*) fruits of thirty-two diverse accessions: variation in health-related compounds and implications for breeding. Phytochemistry. 2011;72(11–12):1358–70. <https://doi.org/10.1016/j.phytochem.2011.03.016>
 15. Meitei AL, Yogendra K, Sanivarapu H, Meitei NT, Rai M, Tyagi W. Metabolite profiling reveals differential accumulation of secondary metabolites related to flavour and colour across four heirloom chilli landraces. Physiol Mol Biol Plants. 2025;31(3):477–91. <https://doi.org/10.1007/s12298-025-01576-4>
 16. Bosland PW, Baral JB. 'Bhut Jolokia'—the world's hottest known chile pepper is a putative naturally occurring interspecific hybrid. HortScience. 2007;42(2):222–4. <https://doi.org/10.21273/HORTSCI.42.2.222>
 17. Mitali B, Minakshi B, Robinson CJ, Rajeev S. Morpho-agronomic variability and multivariate analysis across successive generations of *Capsicum chinense* Jacq. (Bhut Jolokia). Plant Sci Today. 2026. <https://doi.org/10.14719/pst.15448>
 18. Wahyuni Y, Ballester AR, Tikunov Y, de Vos RC, Pelgrom KT, Maharijaya A, et al. Metabolomics and molecular marker analysis to explore pepper (*Capsicum* sp.) biodiversity. Metabolomics. 2013;9(1):130–44. <https://doi.org/10.1007/s11306-012-0432-6>
 19. Xiao JF, Zhou B, Ransom HW. Metabolite identification and quantitation in LC-MS/MS-based metabolomics. Trends Anal Chem. 2012;32:1–14. <https://doi.org/10.1016/j.trac.2011.08.009>
 20. Collins MD, Wasmund LM, Bosland PW. Improved method for quantifying capsaicinoids in *Capsicum* using high-performance liquid chromatography. HortScience. 1995;30(1):137–9. <https://doi.org/10.21273/HORTSCI.30.1.137>
 21. Jaccard P. Étude comparative de la distribution florale dans une portion des Alpes et des Jura. Bull Soc Vaud Sci Nat. 1901;37:547–79.
 22. Jolliffe IT. Principal component analysis. 2nd ed. New York: Springer; 2002.
 23. Kruskal WH, Wallis WA. Use of ranks in one-criterion variance analysis. J Am Stat Assoc. 1952;47(260):583–621. <https://doi.org/10.1080/01621459.1952.10483441>
 24. Dunn OJ. Multiple comparisons using rank sums. Technometrics. 1964;6(3):241–52. <https://doi.org/10.1080/00401706.1964.10490181>
 25. Narasanna R, Mansoori A, Mishra N, Sharma V, Thomas S, Vishwakarma A, et al. Plant metabolomics for crop improvement. In: Agricultural biotechnology: latest research and trends. Singapore: Springer; 2022. p. 385–413. https://doi.org/10.1007/978-981-16-2339-4_17
 26. Aranha BC, Hoffmann JF, Barbieri RL, Rombaldi CV, Chaves FC. Untargeted metabolomic analysis of *Capsicum* spp. by GC-MS. Phytochem Anal. 2017;28(5):439–47. <https://doi.org/10.1002/pca.2692>
 27. Hao Y, Zhang Z, Luo E, Yang J, Wang S. Plant metabolomics: applications and challenges in the era of multi-omics big data. Abiotech. 2025;6(1):116–32. <https://doi.org/10.1007/s42994-024-00194-0>
 28. Heng Z, Xu X, Xu X, Li Y, Wang H, Huang W, et al. Integrated transcriptomic and metabolomic analysis of chili pepper fruits provides new insight into the regulation of the branched-chain esters and capsaicin biosynthesis. Food Res Int. 2023;169:112856. <https://doi.org/10.1016/j.foodres.2023.112856>
 29. Venkatesh J, Lee SY, Back S, Kim TG, Kim GW, Kim JM, et al. Update on the genetic and molecular regulation of the biosynthetic pathways underlying pepper fruit color and pungency. Curr Plant Biol. 2023;35:100303. <https://doi.org/10.1016/j.cpb.2023.100303>
 30. Reddy UK, Karnatam KS, Talavera-Caro A, Lopez-Ortiz C, Ku KM, Chinreddy SR, et al. Uncovering the genetic architecture of pungency, carotenoids, and flavor in *Capsicum chinense* via TWAS-mGWAS integration and spatial transcriptomics. Hortic Res. 2025;12(12):uhaf243. <https://doi.org/10.1093/hr/uhaf243>
 31. He J, Zhang X, Chen S, Yang J, Li Z. Integrated multi-omics approaches provide novel insights into the mechanisms underlying signature flavor development in mulberry fruits. Foods. 2025;14(19):3309. <https://doi.org/10.3390/foods14193309>
 32. Nacius YV, Fernandes GD, Gul K, Ara B, Vieira N, Costa JA, et al. Bioactive potential and phytochemical composition of genus *Capsicum* peppers emphasizing the Amazonian variety Cumari do Pará: a comprehensive review. Pharmacol Res Nat Prod. 2026:100807. <https://doi.org/10.1016/j.prenap.2026.100807>
 33. Arce-Rodríguez ML, Ochoa-Alejo N. An R2R3-MYB transcription factor regulates capsaicinoid biosynthesis. Plant Physiol. 2017;174(3):1359–70. <https://doi.org/10.1104/pp.17.00506>

Additional information

Peer review: Publisher thanks the Sectional Editor and anonymous reviewers for their contribution to the peer review of this work.

Reprints & permissions information: Available at https://horizonpublishing.com/journals/index.php/PST/open_access_policy.

Publisher's Note: Horizon e-Publishing Group remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Indexing: See https://horizonpublishing.com/journals/index.php/PST/indexing_abstracting.

Copyright: © The Author(s). This is an open-access article distributed under the Creative Commons Attribution License CC BY 4.0.

Publisher information: Plant Science Today is published by Horizon e-Publishing Group with support from Empirion Publishers Private Limited, Thiruvananthapuram, India.



Supplementary Table 1. Present and absent matrix on metabolites

Compounds	A303(F1)	A303-06(F2)	A303-06-18(F3)	A425(F1)	A425-10(F2)	A425-10-15(F3)
(2S)-2-[[1-(R)-Carboxyethyl]amino]pentanoate	0	1	0	0	0	0
(2S)-2-hydroxyphytanic acid	0	0	0	1	1	1
(3b,16a,20R)-25-Acetoxy-3,16,20,22-tetrahydroxy-5-cucurbiten-11-one 3-glucoside	1	0	0	0	0	0
(3b,16a,21b,22a)-12-Oleanene-3,16,21,23,28-pentol-22-angeloyloxy-23-al	0	0	0	0	1	1
(9Z,11E,13E,15Z)-4-Oxo-9,11,13,15-octadecatetraenoic acid	1	0	0	0	0	0
(Z)-5-Hexadecenoic acid	1	0	1	1	0	0
10'-apo-beta-carotenal	0	1	1	0	0	0
16-Hydroxy hexadecanoic acid	0	0	0	1	1	0
2,3,4-Trimethyl-3-pentanol	1	1	1	0	0	0
2,3-Diethylpyrazine	1	1	0	0	0	0
23-trans-p-Coumaroyloxytormentic acid	1	1	1	0	0	0
2alpha,17beta-Dihydroxyandrost-4-en-3-one dipropionate	0	0	0	1	1	0
2-Butyl-4-ethyl-5-methyloxazole	0	1	1	0	0	0
2-Hexyl-4,5-dimethyloxazole	1	1	1	0	0	0
2-Hydroxymethylserine	1	1	1	0	0	0
2-Isopropyl-1,4-hexadiene	0	1	0	0	0	0
2-Methyl-2-pentenoic acid	1	0	1	1	1	1
2-Octenedioic acid	1	1	0	0	0	0
2-O-Protocatechuoylaliphilic acid	0	0	0	0	1	1
2S-hydroxy-octadecanoic acid	0	1	0	1	1	1
3-Acetamidopropanal	0	1	0	0	0	0
3beta,6beta-Dihydroxynortropane	0	1	1	0	0	0
3-Hydroxy-b,e-caroten-3'-one	0	0	1	0	0	0
3-hydroxy-tetradecanedioic acid	0	0	0	0	1	1
3-keto stearic acid	0	0	1	0	0	0
4-Ethyl-2-hexyl-5-methyloxazole	0	1	0	0	0	0
4-Hydroxy-3-(16-methylheptadecyl)-2H-pyran-2-one	0	0	0	1	0	1
7,8-Dehydro-beta-micropteroxanthin	1	0	1	0	0	0
8S-HODE	0	0	1	1	1	1
9,10-Dihydroxy-12,13-epoxyoctadecanoate	0	0	0	0	1	1
9-HOTE	0	0	1	0	0	0
9-Hydroxytridecyl docosanoate	1	0	1	1	1	1
Adenine	0	1	1	0	0	0
Adiantifoline	0	0	0	1	1	1
Alloxanthin	1	1	1	0	0	1
all-trans-heptaprenyl diphosphate	1	1	1	1	1	1
alpha-Micropteroxanthin B	1	0	1	0	0	0
Amitraz	0	0	0	1	1	1
Ansatrienin A	1	1	1	0	0	0

Avenestergenin A2	1	1	1	0	0	0
Bismahanine	1	1	1	1	0	0
BQ 123	1	1	1	1	1	1
Caffeoylcycloartenol	1	1	1	0	0	0
Capreomycin	0	0	0	1	1	1
Capsi-amide	1	1	1	0	0	0
Capsoside A	1	1	1	1	1	1
Cavipetin D	0	0	0	0	1	0
Columbianetin	0	1	0	0	0	0
Dalfopristin	0	0	0	1	1	1
Delavaine A	1	1	1	0	0	0
Elatine	1	1	1	0	0	0
Feruloyl-beta-sitosterol	1	0	0	0	0	0
Feruloyldihydro-beta-sitosterol	1	1	1	0	0	0
Gabunamine	0	1	1	0	0	0
Ganoderic acid Md	1	0	1	0	0	0
Glycerol 5-hydroxydecanoate	0	0	0	0	1	1
Gnididilatin	1	1	1	1	1	1
Homocapsaicin	1	1	1	0	0	0
Homodihydrocapsaicin	1	1	1	0	0	0
Juvenile hormone III	0	1	0	0	0	0
Lansioside C	1	0	1	1	1	1
Linalyl caprylate	1	1	1	1	1	1
L-trans-5-Hydroxy-2-piperidinecarboxylic acid	1	1	1	0	0	0
Lyciumoside III	1	1	1	1	0	0
Marshdimerin	0	0	0	1	1	0
Melilotoside A	0	0	0	1	1	0
Methyl 2-furoate	0	1	0	0	0	0
Momordicoside E	0	0	1	0	0	0
Mosinone A	1	0	0	0	1	0
N,2,3-Trimethyl-2-(1-methylethyl)butanamide	1	1	1	1	1	1
N-[(Ethoxycarbonyl)methyl]-p-menthane-3-carboxamide	1	0	1	0	0	0
N1,N5,N10-Tris-trans-p-coumaroylspermine	1	1	1	1	0	1
NAC-FnorLRF-amide	1	0	0	0	0	0
Nitramine	1	1	1	0	0	0
Nordihydrocapsaicin	1	1	1	0	0	0
Nummularine F	1	1	0	0	0	0
Octyl octanoate	0	0	0	1	1	1
omega-hydroxy behenic	0	0	0	1	0	0
Palmitic Acid	1	0	1	0	0	0
Palmitic amide	1	1	1	0	0	0
Pergolide	0	1	1	0	0	0
Petroselinic acid	1	0	0	1	1	1
PGD2-d4	1	0	1	0	0	0

Phenylacetic acid	1	1	1	0	0	0
Pimelea factor P2	1	1	1	1	1	1
Retronecine	0	1	0	0	0	0
Rhodoxanthin	1	1	1	1	1	1
Sanguisorbin B	0	0	0	1	1	1
Saringosterol 3-glucoside	0	0	0	1	0	0
Sibutramine	0	0	0	1	1	1
Spinoside A	0	0	0	1	1	0
Staphidine	0	0	0	1	1	1
Stearaldehyde	0	0	0	1	0	1
Tecostanine	1	1	1	0	0	0
Tetrahydropteridine	0	0	1	0	0	0
Tetralin	1	0	0	0	0	0
Thalicarpine	0	0	0	0	1	1
Toluate	1	1	1	0	0	0
Tridemorph	1	1	1	0	0	0
Trolamine	0	1	0	0	0	0
Voacamine	1	1	1	0	0	1
xi-7-Octen-2-ol	1	0	0	0	0	0

F1, F2 and F3: first, second and third selfed generations, respectively; LC–MS/MS: liquid chromatography–tandem mass spectrometry; HPLC: high-performance liquid chromatography; SHU: Scoville heat units. Sample identification key: For LC–MS/MS metabolomic profiling, A303, A303-06, and A303-06-18 represent the F1, F2, and F3 generations, respectively, of the A3 lineage. Similarly, A425, A425-10, and A425-10-15 represent the F1, F2, and F3 generations, respectively, of the A4 lineage. The HPLC data are presented using individual plant-level sample IDs, grouped by generation (F1, F2, and F3) for SHU analysis.

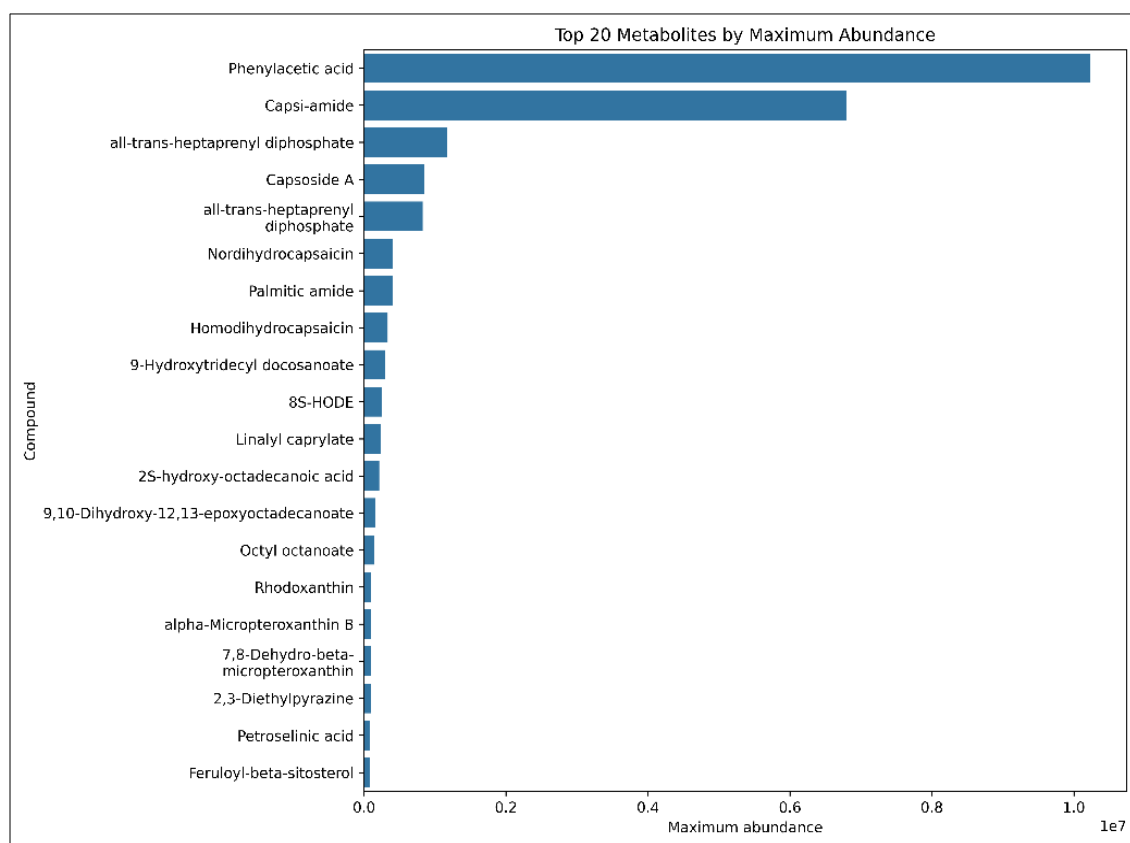
Supplementary Table 2. Peak areas of capsaicin and dihydrocapsaicin and Scoville heat units (SHU) across successive selfed generations

Generation	Sample id	peak area cap	peak area dc	SHU
F1	A303	13727314	4300424	907456
F1	A312	9032917	2768795	594498
F1	A313	11438884	3537379	754187
F2	A30304	41760036	81841941	1679935
F2	A30306	35012367	11807531	689663.1
F2	A30334	3409753	11443087	197404.8
F2	A31202	37199630	12516950	732390.2
F2	A31215	33511746	11375848	661045.6
F2	A31220	37077041	12430293	729402.1
F2	A31333	35074575	11745691	689840.6
F2	A31334	33605073	11433593	663216.3
F2	A31336	32834077	10977110	645543.5
F3	A3030403	4045704	1931118	78180.74
F3	A3030413	4171815	2016373	80903.15
F3	A3030434	4050767	1953973	78511.34
F3	A3030612	4446974	2137710	86106.5
F3	A3030618	7591695	3653688	147046.2
F3	A3030622	4130804	2012370	80287.98
F3	A3033410	5819882	2873313	113551.4

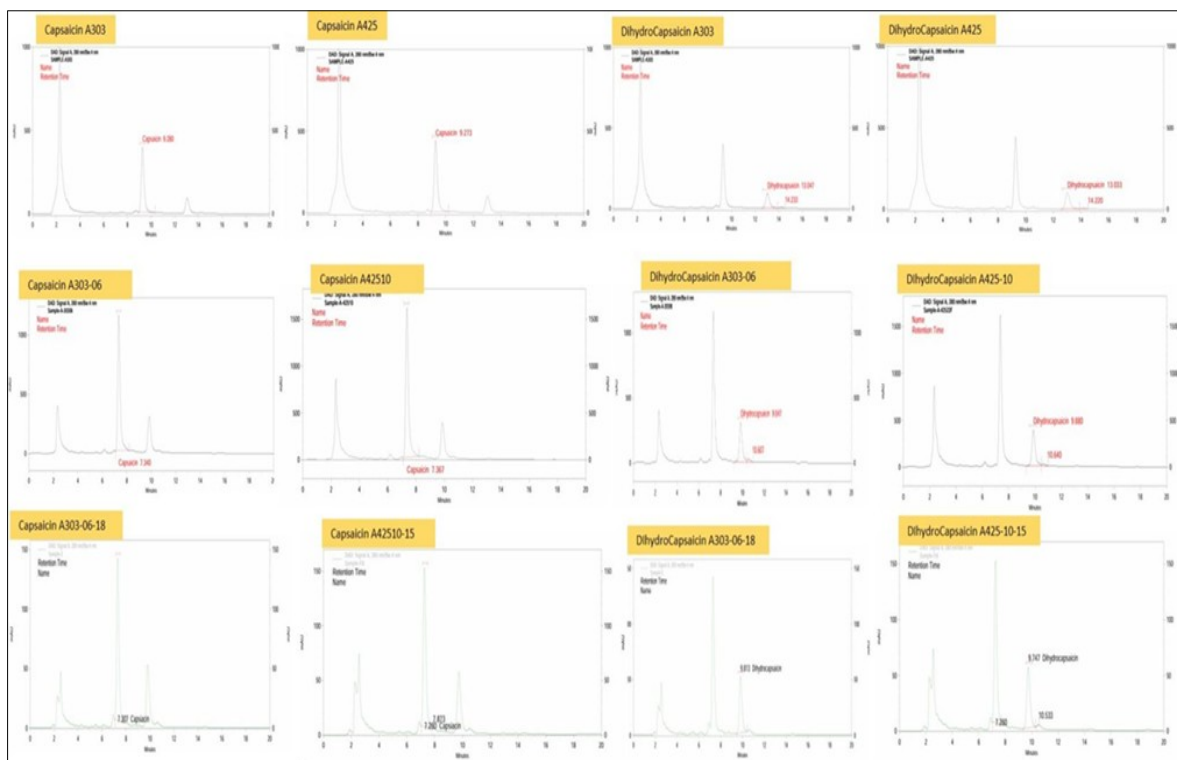
F3	A3033413	4025991	1951280	78136.58
F3	A3033416	3964686	1936739	77119.55
F3	A3120207	3980360	1972312	77742.36
F3	A3120224	5205027	2587802	101760.5
F3	A3120228	5128222	2555748	100328.7
F3	A3121502	7159776	3540243	139755.6
F3	A3121510	3941462	3725985	97173.78
F3	A3121518	6910197	3455723	135326.8
F3	A3122016	5026967	2558796	98957.19
F3	A3122020	5275586	2607729	102967.4
F3	A3122021	5206467	2546938	101315.1
F3	A3133416	3985143	1923938	77257.88
F3	A3133422	4097730	1976664	79422
F3	A3133424	4127027	1994905	80036.63
F3	A3133304	7630697	1912887	127762.9
F3	A3133315	7364200	3520667	142371.8
F3	A3133316	7537879	3638576	146126.7
F3	A3133609	3998262	1937992	77600.14
F3	A3133615	4030037	1969140	78396.17
F3	A3133623	5101169	2521132	99558.8
F1	A401	8547549	2627116	562860
F1	A406	6478730	5073943	559677
F1	A425	15388632	4752482	1014329
F2	A40117	32642452	11141691	644669
F2	A40118	35920508	12172598	708295.5
F2	A40119	46096194	15804313	911264.7
F2	A40601	34163034	11477736	672384.5
F2	A40602	42170044	14176053	830079.4
F2	A40603	36545273	12302014	719573.5
F2	A42502	35478016	12037620	699759.7
F2	A42510	48420139	13220847	914435.8
F2	A42522	39169779	14656562	789863.5
F3	A4011702	6102825	2977524	118667.8
F3	A4011705	4649681	2279662	90538.39
F3	A4011710	4417886	2171441	86086.66
F3	A4011806	5921540	2881291	115054.1
F3	A4011822	7213235	3367950	138535.9
F3	A4011828	5970702	2929957	116291.1
F3	A4011910	5561242	2728021	108304.6
F3	A4011915	5717638	2786142	111138.6
F3	A4011925	4376440	2147550	85238.96
F3	A4060115	4565897	2251164	89050.22
F3	A4060118	6710322	3271059	130447.8
F3	A4060120	4674364	2319252	91332.06
F3	A4060202	4228566	3544844	99098.27

F3	A4060210	6419250	3145088	124970.7
F3	A4060212	4708143	2343203	92073.96
F3	A4060308	5689345	2787289	110758.7
F3	A4060310	5441486	2685798	106160.5
F3	A4060322	4641105	2289992	90536.92
F3	A4250205	6836834	2067362	118496.6
F3	A4250210	4132483	2045068	80683.68
F3	A4250215	4866285	2435431	95320.64
F3	A4251005	4219169	2081547	82303.05
F3	A4251018	12314298	6016866	239548.5
F3	A4251027	4504652	2193891	87547.37
F3	A4252208	4488870	2159847	86940.47
F3	A4252215	5530059	2645592	106932.8
F3	A4252232	3941404	1886573	76224.89

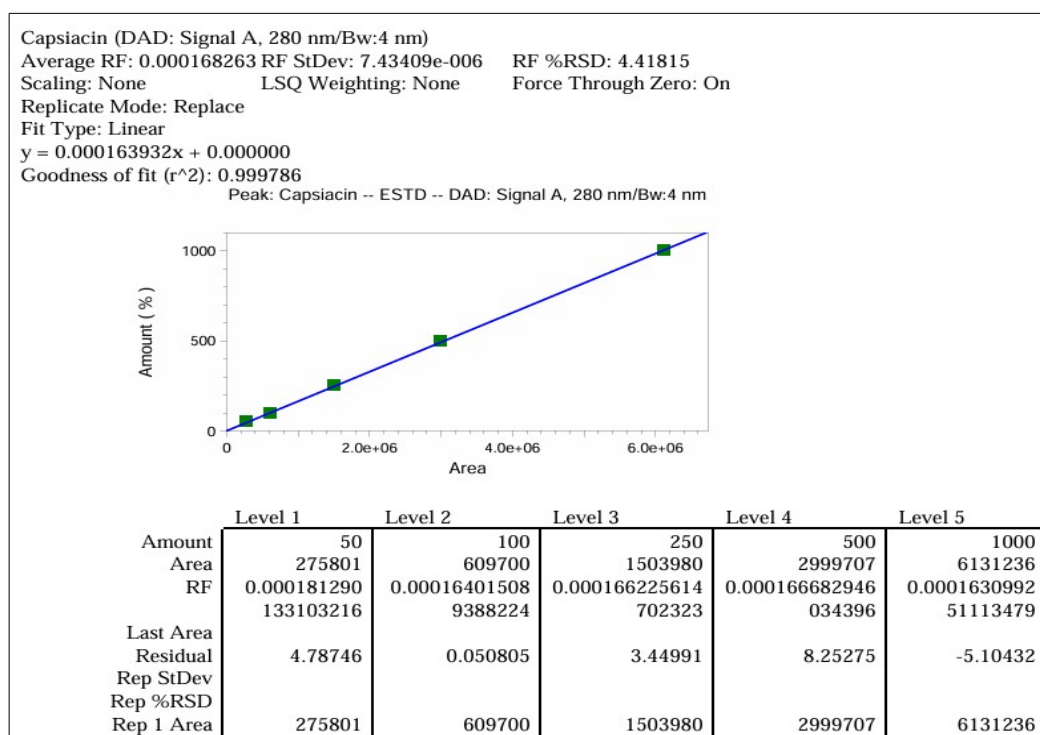
F1, F2 and F3: first, second and third generations, respectively; cap: capsaicin; DC: dihydrocapsaicin; SHU: Scoville heat units; ppm: parts per million. Peak area cap: HPLC peak area of capsaicin; peak area DC: HPLC peak area of dihydrocapsaicin; SHU: Scoville heat units calculated from the total capsaicinoid concentration (ppm) following the ASTA analytical method



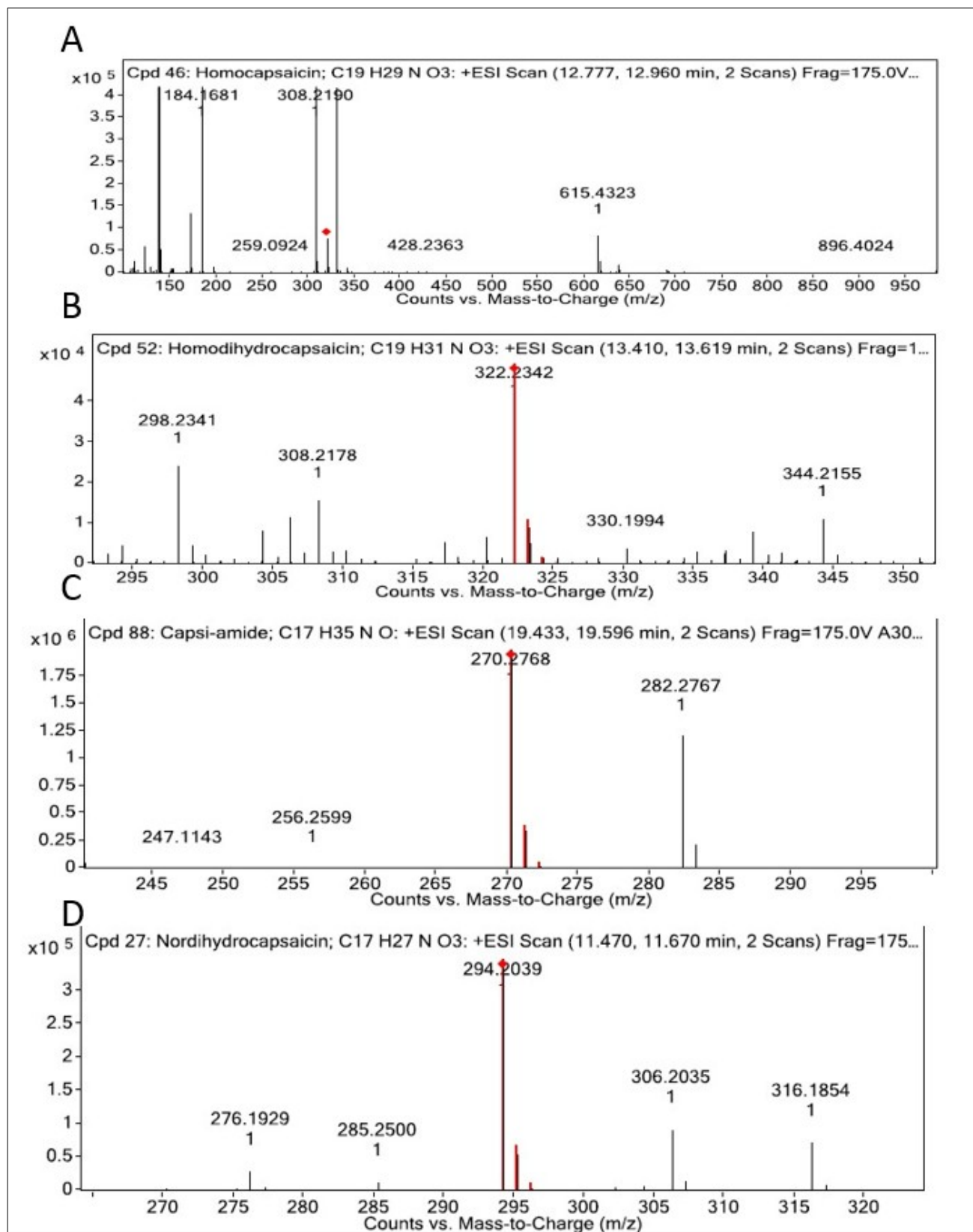
Supplementary Fig. 1. Mass accuracy and database support of the putatively annotated metabolites, categorised according to annotation quality (tentative, moderate, and high).



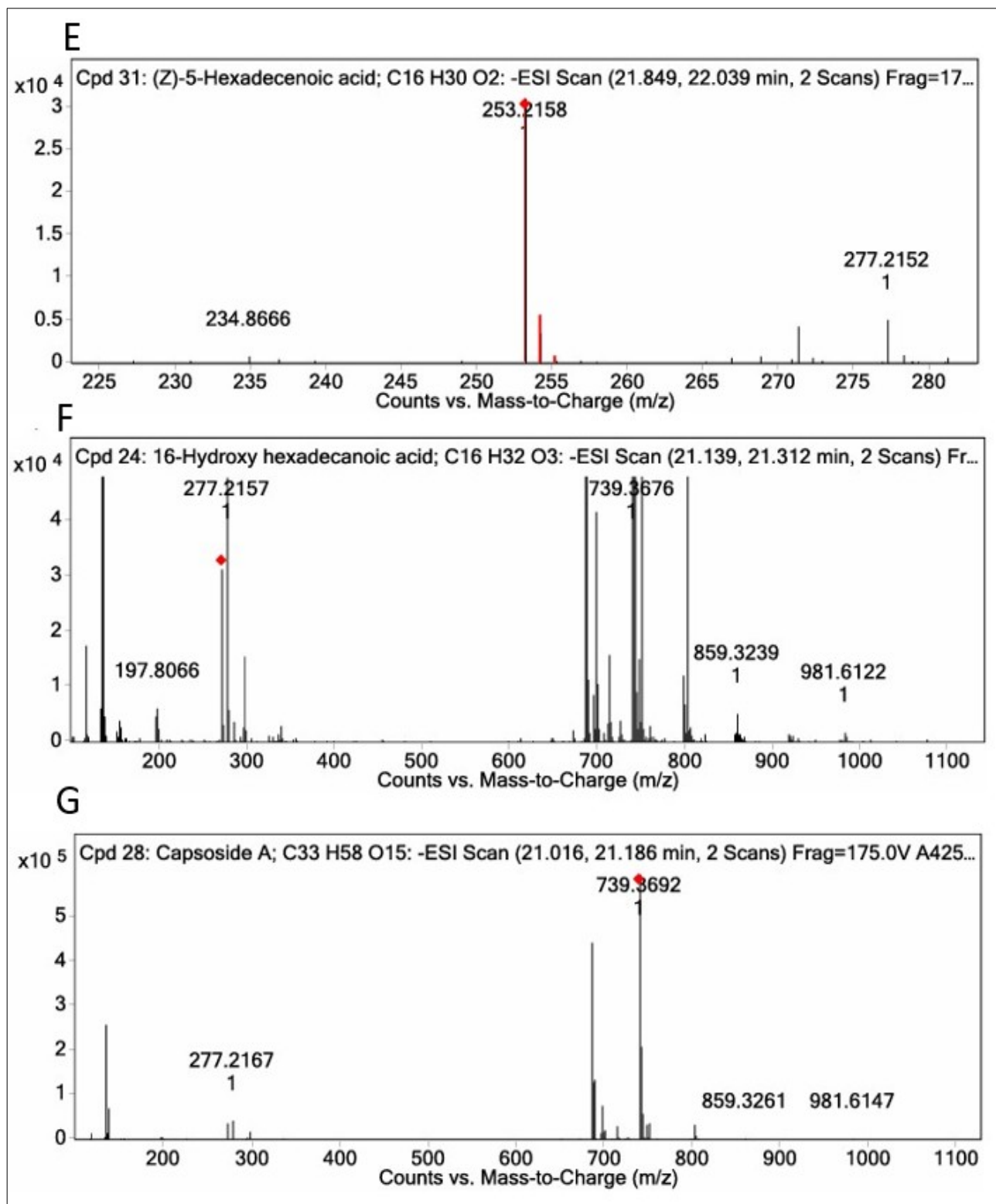
Supplementary Fig. 2. Representative high-performance liquid chromatography (HPLC) chromatograms of *Capsicum chinense* fruit samples from the A3 lineage (A303, A303-06, and A303-06-18) and the A4 lineage (A425, A425-10, and A425-10-15), showing the characteristic peaks of capsaicin and dihydrocapsaicin. The chromatograms were recorded at 280 nm and were used for the identification and quantification of the major capsaicinoids.



Supplementary Fig. 3. Calibration curve of the capsaicin standard showing the linear relationship between peak area and concentration (50–1000 ppm) determined by high-performance liquid chromatography–diode array detection (HPLC–DAD) at 280 nm. The calibration exhibited excellent linearity ($R^2 = 0.999786$) and was used to quantify capsaicin in *Capsicum chinense* samples.



Supplementary Fig. 4. Liquid chromatography–tandem mass spectrometry (LC–MS/MS) spectra of important compounds detected in the A3 lineage. Note: (A) Homocapsaicin; (B) Homodihydrocapsaicin; (C) Capsi-amide; (D) Nordihydrocapsaicin.



Supplementary Fig. 5. Liquid chromatography–tandem mass spectrometry (LC–MS/MS) spectra of some important compounds detected in the A4 lineage. (A) (Z)-5-Hexadecenoic acid; (B) 16-Hydroxy hexadecanoic acid; (C) CapsosideA.