Article:

Cawood, G.L. and Ton, J. (2025) Decoding resilience: ecology, regulation, and evolution of biosynthetic gene clusters. Trends in Plant Science, 30 (2). pp. 185-198. ISSN: 1360-1385

https://doi.org/10.1016/j.tplants.2024.09.008

Trends in
Plant Science

CellPress

Review

Decoding resilience: ecology, regulation, and evolution of biosynthetic gene clusters

George Lister Cawood[1,*] and Jurriaan Ton[1,*]

Secondary metabolism is crucial for plant survival and can generate chemistry with nutritional, therapeutic, and industrial value. Biosynthetic genes of selected secondary metabolites cluster within localised chromosomal regions. The arrangement of these biosynthetic gene clusters (BGCs) challenges the long-held model of random gene order in eukaryotes, raising questions about their regulation, ecological significance, and evolution. In this review, we address these questions by exploring the contribution of BGCs to ecologically relevant plant–biotic interactions, while also evaluating the molecular-(epi)genetic mechanisms controlling their coordinated stress- and tissue-specific expression. Based on evidence that BGCs have distinct chromatin signatures and are enriched with transposable elements (TEs), we integrate emerging hypotheses into an updated evolutionary model emphasising how stress-induced epigenetic processes have shaped BGC formation.

Highlights

The discovery of biosynthetic gene clusters (BGCs) in plants has stimulated crossdisciplinary research into their ecological significance, regulation, and evolution.

BGCs reside in dynamic genome regions enriched with transposons, histone variant H2A.Z, and histone modification H3K27me3. These configurations enable tightly coordinated stress- and tissue-dependent expression, while also facilitating genomic plasticity and rapid adaptation to stressful environments.

Negative selection against deleterious pathway fragmentation, combined with positive selection for beneficial functions in defence and ecological interactions, preserves genetic linkage and BGC integrity.

Future research into the molecular regulation and ecological significance of BGCs will shed light on the catalytic influences of stress-inducible epigenetic changes and transposons in the evolution of plant secondary metabolism.

Role of biosynthetic gene clusters in plant metabolism

The number of low-molecular-weight natural products in the plant kingdom amounts up to 1 million chemical structures [1], which includes economically important molecules with nutritional, industrial, and therapeutic value [2]. Primary metabolites, such as (poly)saccharides, amino acids, and organic acids, provide structure and energy to ensure physiological functioning through cell division, growth, and reproduction. These metabolites are conserved within the plant kingdom [3] and mutations in the underpinning genes are constrained to conserve stable functioning [4]. By contrast, secondary metabolites (now more commonly referred to as specialised metabolites), such as flavonoids, terpenoids, and nitrogen and sulfur-containing compounds, are far more diverse and tend to be lineage specific [5,6].

The stress hormones salicylic acid (SA), jasmonic acid (JA), and abscisic acid (ABA) contribute to the regulation of secondary metabolites with roles in plant defence [7]. Most secondary metabolites are derived from products of primary metabolic pathways, such as the shikimate pathway for aromatic amino acid biosynthesis [8], or from **neofunctionalisation** (see Glossary) of enzymes in these pathways [9]. Some defence-related secondary metabolites, such as phytoanticipins, are constitutively produced in specific tissues [10], while others, such as phytoalexins, are synthesised *de novo* in response to environmental stress [4]. Due to the costs involved, biosynthesis of these metabolites is tightly controlled in space and time, enabling production within specialised cell types during the more vulnerable growth stages and/or exposure to stress [11,12]. Apart from their direct antifeeding and biocidal activities against pest and diseases [13], secondary metabolites also mediate beneficial ecological interactions, such as the recruitment of pollinators and seed dispersers [14] or intraspecific plant–plant communication [15,16]. Secondary metabolites can also act as internal defence signals, for instance by controlling the induction of early postinvasive defences against pathogens and aphids [17]. However, the strict distinction between primary and secondary metabolism in plants is not clearcut, with some secondary metabolites

[1] Plants, Photosynthesis and Soil, School of Biosciences, The University of Sheffield, Western Bank, Sheffield, S10 2TN, UK

*Correspondence:
gdlcawood1@sheffield.ac.uk
(G.L. Cawood) and j.ton@sheffield.ac.uk
(J. Ton).

acting as precursors to primary metabolites or having other regulatory functions over primary metabolism [18].

The pathways controlling plant secondary metabolites involve complex networks of enzymes, generating chemistry with diverse biological activities [4]. While most biosynthetic genes are distributed randomly across the genome, examples of specialised metabolic products are encoded by biosynthetic genes which cluster together within localised chromosomal regions [19]. These so-called BGCs are found in both plants and fungi and challenge the long-held model of random gene organisation in eukaryotes [20]. BGCs comprise at least three co-expressed nonhomologous genes [21], which are spatially and functionally linked to mediate tightly controlled synthesis of secondary metabolites. Recent years has seen a rapidly expanding research field focussing on explaining the emergence and maintenance of these highly improbable genetic groupings.

Here, we review the latest insights into the contribution of BGCs to plant–biotic interactions and critically discuss the benefits of clustering in terms of regulation by genetic and epigenetic mechanisms. We furthermore draw parallels with the genomic clustering of immune-related genes, integrating and refining emerging hypotheses to present a consolidated model that highlights a central role for stress-induced epigenetic processes in BGC formation, and the antagonistic selective pressures preserving BGC integrity.

The defence symphony: biosynthetic gene clusters in direct defence and ecological interactions to protect plants

Arabidopsis (*Arabidopsis thaliana*) contains four BGCs producing tri-terpenes. The arabidiol cluster is the largest and harbours 11 genes that span an estimated 83 kbp on chromosome 4 [22]. The other three tri-terpene clusters are located on chromosome 5 and include the well-studied thalianol cluster, which comprises four co-expressed genes clustered across a 56-kbp region [23], *THAA2,* which is separated by two intervening genes in accession Col-0 [24], and two satellite genes, *THAR1* and *THAR2*, on chromosomes 3 and 1, respectively (Figure 1). The BGCs controlling thalianol, arabidiol, and marneral form an interconnected network that is modulated by promiscuous acetyl-transferases and alcohol dehydrogenases to produce >50 root-secreted compounds, derivatives of which shape the rhizosphere microbiome [22]. The curation of these arabidopsis-specific microbial communities likely provides benefits through enhanced nutrient acquisition and disease suppression [25]. Tri-terpenes also have roles in direct defensive interactions, particularly in the rhizosphere. For instance, the oxidative degradation of root-produced arabidiol into the homoterpene (*E*)-4,8-dimethyl-1,3,7-nonatriene (DMNT) controls infection by the root-rot pathogen *Pythium irregulare* [26].

Another well-studied plant BGC is responsible for the production of indole-derived benzoxazinoids in members of the Poaceae. In maize, the most dominant benzoxazinoids are 2,4-dihydroxy-7-methoxy-1,4-benzoxazin-3-one glucoside (DIMBOA-glc) and 2-hydroxy-4,7-dimethoxy-1,4-benzoxazin-3-one glucoside (HDMBOA-glc). The first eight biosynthetic enzymes toward DIMBOA-glc and HDMBOA-glc are encoded by a core cluster of *Bx* genes on the short arm of chromosome 4 [27] and a satellite gene on chromosome 1 (Figure 2). Deglucosylated benzoxazinoids, such as DIMBOA and HDMBOA, have broad-spectrum biological activities, including antifeedant, antibiotic, and defence-signalling properties, which provide defence against chewing insects, phloem-feeding herbivores, and pathogenic fungi [28,29]. Root-exuded DIMBOA also acts as a **semiochemical** in the rhizosphere, recruiting *Pseudomonas putida* KT2440 [30], which primes JA-controlled indirect defences [31]. More recent studies have shown that exudation of benzoxazinoids can alter bacterial and fungal communities in the soil and rhizosphere, providing long-term benefits in the form of pest resistance that can even benefit

Glossary

Chromatin topology: 3D spatial arrangement of chromatin, including compaction, loop formation, and long-range interactions.
Ectopic recombination: recombination of homologous DNA sequences at nonallelic regions in distal genomic locations; implicated in genome rearrangements and gene duplication.
Euchromatin: loosely packed chromatin associated with a transcriptionally responsive/active state of gene expression.
Exon shuffling: formation of new genes through recombination of intron sequences, resulting in duplications and rearrangements of existing exons.
Gene fusion: fusion of two independent gene sequences to produce a hybrid gene, caused by structural genomic rearrangements, such as inversions, translocations, and duplications.
Genetic linkage: proximal genes with a high probability of being inherited together and a low probability of being uncoupled through recombination.
Heterochromatin: densely packed chromatin associated with a transcriptionally silenced state of gene expression.
Histones: proteins that interact with DNA to produce nucleosomes. Post-translational modifications alter their interaction with DNA and nuclear proteins, promoting or repressing transcription.
Long noncoding (lnc)RNA: RNAs exceeding 200 nucleotides that are not translated into functional proteins; involved in gene regulation and chromatin architecture, either directly, or by affecting the activity of regulatory proteins.
Monocistronic: mRNA that is transcribed from a single cistron and translated into a single polypeptide/protein.
Neofunctionalisation: acquisition of a new adaptive function by a duplicated gene through mutation(s) in coding and/or *cis*-regulatory regions.
Nonallelic homologous recombination (NAHR): see 'Ectopic recombination'.
Pericentromere: regions on the adjacent sides of the chromosomal centromere; enriched with transposons and heterochromatic DNA.
Polycistronic: mRNA from a prokaryotic operon comprising more

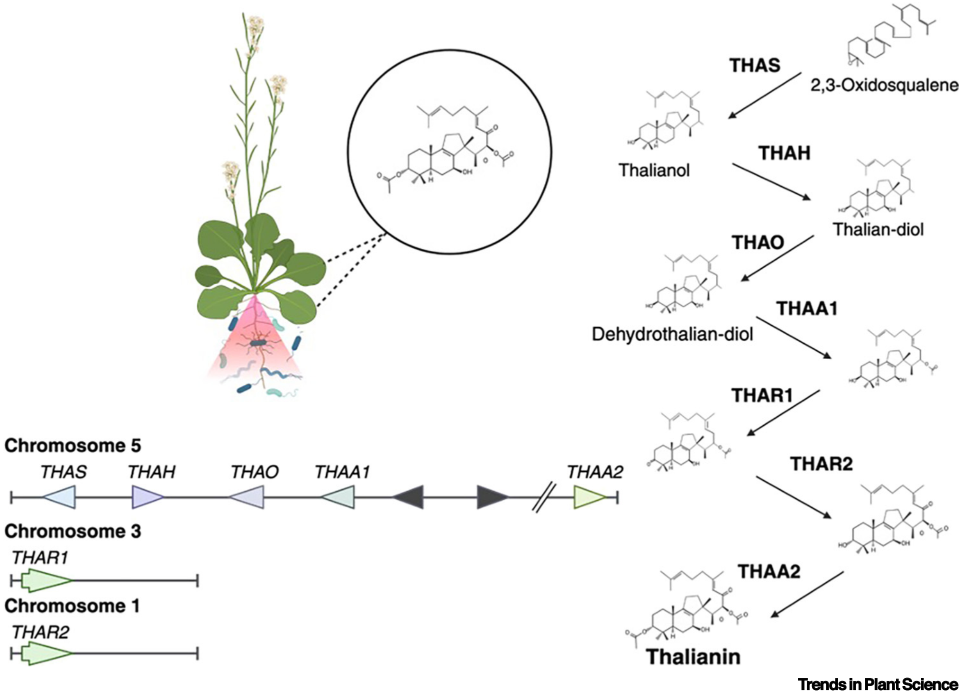


Figure 1. Tri-terpene biosynthesis in arabidopsis (*Arabidopsis thaliana*). The thalianol biosynthetic gene cluster (BGC) encoding the production of the tri-terpene thalianol, and derivatives thereof, has a role in rhizosphere interactions through the recruitment and/or selection of bacterial taxa that may promote plant growth and health [22,25]. The oxidosqualene cyclase thalianol synthase (THAS) converts 2,3-oxidosqualene into thalianol, which is then converted to thalian-diol and dehydrothalian-diol by two cytochrome P450s, THAH and THAO, followed by acetylation through the *THAA1*-encoded acyltransferase. The acetyltransferase-encoding *THAA2*, which is separated from the first four cluster elements by two intervening genes in Col-0, but is directly next to them in other arabidopsis accessions, along with the auxiliary oxidoreductases *THAR1* and *THAR2* located on chromosomes 3 and 1, respectively, catalyse subsequent chemical conversions into thalianin. Biosynthetic pathway adapted from [24]. Figure created with BioRender (biorender.com).

plants of the next generation [32–35]. Beyond its contribution to plant defence, DIMBOA has been co-opted by attackers. For instance, the iron-chelating activity of DIMBOA as a phytosiderophore is exploited by DIMBOA-tolerant larvae of *Diabrotica virgifera virgifera* to localise crown roots of maize for feeding [36], highlighting the complex impacts of BGC-encoded secondary metabolites on ecological interactions.

In tomato, BGCs control the production of specific steroidal glycoalkaloids (SGAs) and acyl sugars. The α-tomatine BGC includes six genes that are contiguously clustered across a ~200-kbp region on chromosome 7 [37], with satellite subgroups on chromosome 12 [38]. Broadly classified as a saponin, α-tomatine has been reported to shape the rhizosphere microbiome through enrichment with growth-promoting and resistance-inducing bacterial taxa [39]. Furthermore, α-tomatine has a role in defence as an antifeeding agent against herbivores [40] and inhibition of microbial pathogens through cell necrotisation [41]. However, like benzoxazinoids, selected pathogens have coevolved to detoxify α-tomatine: the necrotrophic fungus *Botrytis cinerea* deglucosylates α-tomatine into the less toxic derivatives β_1-tomatine and tetrasaccharide lycotetraose, inhibiting inducible defences in tomato [42,43]. The acyl-sugar controlling BGC in tomato is located on chromosome 7, with a satellite region on chromosome 12 [44]. It was previously thought that this BGC was only active in glandular trichomes, but recent evidence showed that it is also expressed in root cells [45]. Interestingly, the *SlASAT1-LIKE*

than one cistron; features multiple translational start sites and encodes more than one protein.

Retroduplication: gene duplication by reverse transcription of pre-mRNA and subsequent reintegration at new genomic loci; precedes post-transcriptional modifications of pre-mRNA, retaining introns and untranslated regions (UTRs).

Retroposition: insertion of repetitive DNA sequences into the genome via transcription, RNA processing, reverse transcription, and reintegration via nonhomologous end-joining. Unlike **retroduplication**, the insertion site is not duplicated, and introns and UTRs are lost, resulting in pseudogenes.

Retrotransposition: mobilisation and insertion of retrotransposons. A TE-encoded reverse transcriptase reverse transcribes the RNA intermediate, and the resulting cDNA is inserted into the genome by a TE-encoded integrase.

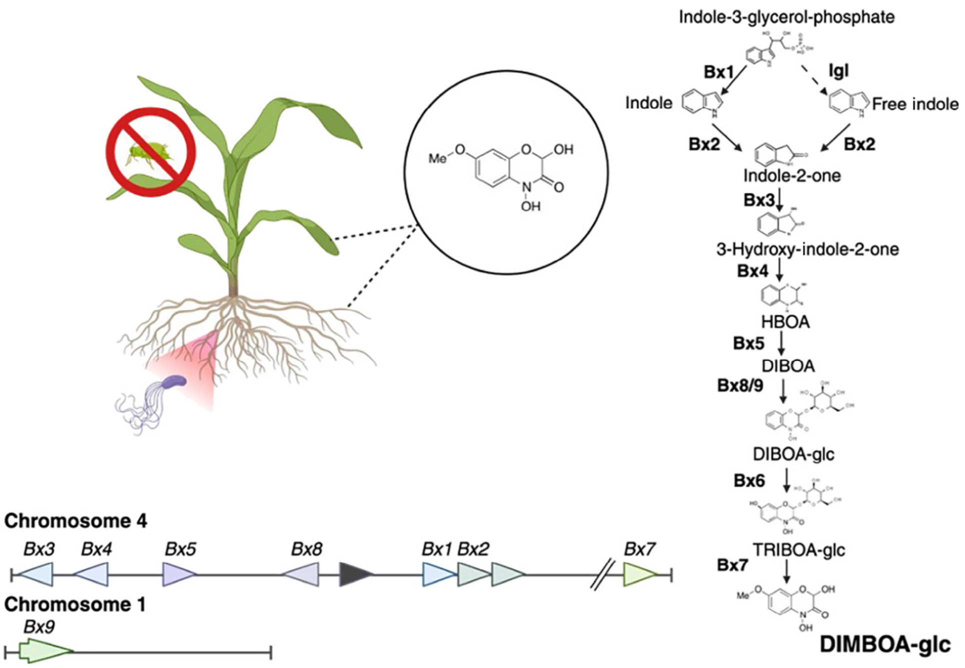
Semiochemical: chemical compound that can alter the behaviour of other organisms upon exposure.

Subfunctionalisation: sharing of the partial functionality of a progenitor gene by a duplicated gene.

Super-enhancer (SE): genomic regions containing multiple enhancers, either upstream or downstream from the genes they regulate. Transcription of distant genes is promoted through loop formation and transcription factor recruitment.

Topologically associated domains (TADs): genomic regions that have enhanced associations with each other through chromatin-loop formation, influencing the transcriptional status of associated genes.

Transduplication: acquisition of fragments from protein-coding genes during TE transposition, leading to the production of novel protein-coding sequences following insertion.



Trends in Plant Science

Figure 2. Benzoxazinoid biosynthesis in *Zea mays*. The benzoxazinoid biosynthetic gene cluster (BGC) of maize controls the production of 2,4-dihydroxy-7-methoxy-1,4-benzoxazin-3-one (DIMBOA) and derivatives thereof, which have diverse roles in defence-related plant–biotic interactions [28–35]. The indole-3-glycerolphosphate lyase enzymes BX1 and IGL convert the pathway precursor into indole, followed by stepwise hydroxylations by four P450 monooxygenases, Bx2–5, resulting in DIBOA [27]. Two UDP-glucosyltransferases, encoded by *Bx8* and the satellite gene *Bx9*, glycosylate DIBOA before subsequent hydrolysation into TRIBOA-glucoside (glc) by the 2-oxoglutarate-dependent dioxygenase Bx6. Methylation of TRIBOA-glc by a methyltransferase encoded by the peripherally located *Bx7* gene produces DIMBOA-glc. Biosynthetic pathway graphic adapted from [32]. Figure created with BioRender (biorender.com).

(*SlASAT1-L*) paralog within the acyl-sugar BGC is required for root-specific biosynthesis of acyl-sugars with glucosylinositol disaccharide cores, which differ from the BGC-produced tri- and tetra-acylsucroses in trichomes [46]. The stress responsiveness and implications for defence and ecological interactions of root produced acyl-sugars are yet to be characterised, but offer the potential to study the regulation and ecological relevance of clustered metabolic pathways.

To summarise, BGC-encoded secondary metabolites have important roles in defence-related ecological interactions, and their efficacy is under constant selective pressure from coevolving aggressors aiming to subvert these defensive functions [42,47]. Notably, there are ample other defence-related secondary metabolites that are not produced by BGCs, implying that BGCs are the result of specific evolutionary mechanisms and pressures. Recent insights into the physiological and ecological costs and benefits of BGCs, as well as their structural genomic characteristics, are beginning to offer plausible explanations as to how and why BGCs have evolved in plants.

Better together: multilayered regulation in space and time

Tight control of secondary metabolite production is essential to minimise allocation costs, storage constraints [23,48], self-toxicity, and coevolution by specialised attackers [49]. Synchronous, rapid, and tightly controlled production is especially important for phytoalexins to ensure coordinated and cost-efficient defence against localised and transient attacks by pests or diseases.

Here, we detail how the clustered nature of BGCs promotes tightly coordinated co-expression in response to environmental stresses, acting as relay points that facilitate synchronous integration of external stress stimuli with expeditious control over the biosynthesis of environmental signals mediating both direct and indirect defence. In prokaryotes, BGCs occur as operons, which are transcribed into a single **polycistronic** pre-mRNA, whereby transcription is regulated by repressor/activator proteins that directly interact with the products of the pathway in a negative feedback loop [50]. By contrast, eukaryotic BGCs are transcribed as **monocistronic** units [50,51], and there is no current evidence to suggest that regulatory factors directly interact with the pathway products. Here, the concerted activities of phytohormones, transcription factors (TF) networks, and epigenetic mechanisms mediate tightly controlled co-expression of BGCs to ensure cost-efficient and spatiotemporal delivery of defence compounds and semiochemicals.

Regulation by transcription factors and phytohormones

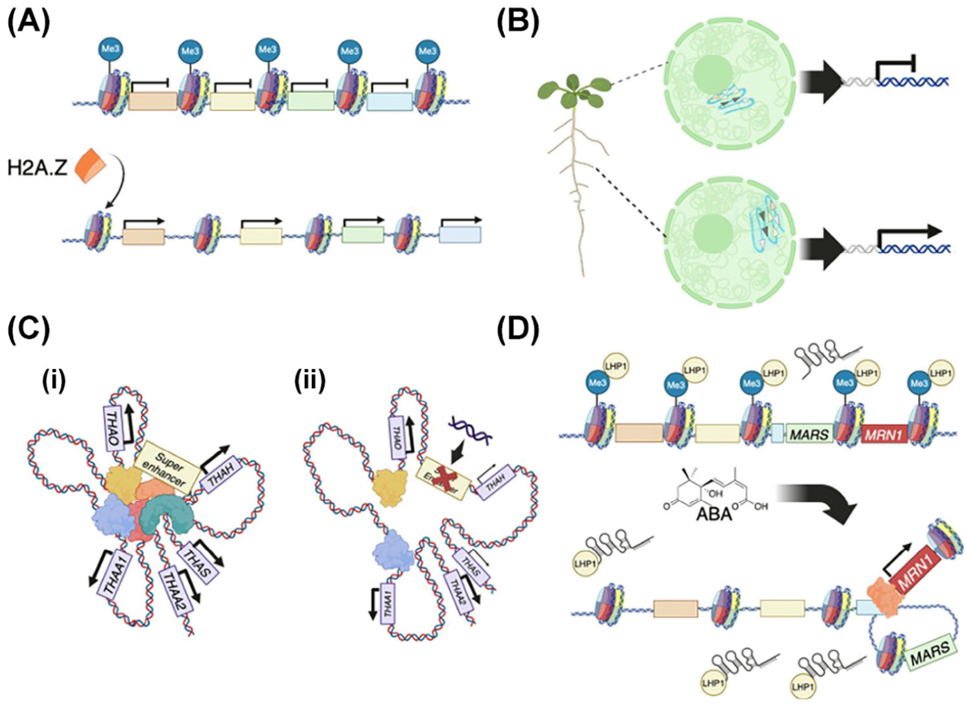
Jasmonates are stress-responsive phytohormones with a vital role in the regulation of secondary metabolism and immunity against necrotrophic pathogens and chewing herbivores [52]. In arabidopsis, stress-induced accumulation of JA-isoleucine induces proteasomal degradation of Jasmonate Zip Domain (JAZ) repressor proteins by the F-Box protein Coronatine Insensitive1 (COI1), which liberates the basic helix-loop-helix TFs MYC1 and MYC2 to induce JA-dependent defence genes [53]. Bai *et al.* [54] showed that loss-of-function mutants in NOVEL INTERACTOR OF JAZ (NINJA), which represses MYC-dependent transcription of JA-inducible genes [55], displays enhanced accumulation of thalianin and its derivatives. Moreover, a combination of single cell RNA-sequencing, mutant profiling, and transient promoter transactivation assays revealed that JA, COI1, and MYC2-4 act in conjunction with bHLH class IVa and homeodomain TFs to mediate cell-specific and stress-inducible expression of the marneral and thalianol BGCs [56]. Both classes of tri-terpene were strongly upregulated in the outer cell layers of root tips in response to methyl-JA treatment, which was dramatically reduced in the *coi1-1* mutant and partially inhibited in the *myc234* triple mutant. Furthermore, chromatin immunoprecipitation (ChIP) sequencing confirmed that MYC2 and MYC3 bind thalianol and marneral BGC gene promoters, with transactivation assays confirming MYC-dependent activation [56]. Interestingly, it was also found that the TF DAG1 inhibits MYC-dependent expression in internal root cells by transactivation repression, explaining why their JA-inducible expression is limited to the outer cell layers of the root tip [56]. In tomato, components of the JA pathway have also been implicated in the regulation of BGC-encoded SGAs. For instance, *myc1 myc2* hairy root cultures displayed reduced production of α-tomatine and dehydrotomatine [57]. Interestingly, however, the same study revealed that CRISPR-cas9 mutagenesis of MYC-binding promoter G-box elements in SGA precursor genes only affected basal transcription and not JA-induced expression, suggesting that the responsiveness of SGAs to stress involves additional regulatory layers.

Although BGC expression involves regulation by phytohormones and defence-related TFs, it remains unclear whether these signalling components were associated with a single cluster component before duplication, neofunctionalisation, and cluster expansion, or whether their recruitment succeeded cluster formation to mitigate the costs of inappropriate overexpression. Interestingly, bioinformatically predicted BGCs, which are physically clustered but do not generate associated products, do not show co-expression [58], suggesting that clustering alone is not responsible for co-expression. However, direct experimental evidence for the benefits of BGCs compared with nonclustered biosynthetic pathways is lacking, and further research is needed to determine whether TF-mediated regulation is sufficient for synchronised BGC expression, or whether other regulatory mechanisms are needed.

Cis-regulation by histone modifications and variants

Apart from phytohormones and TFs, epigenetic mechanisms, such as post-translational modifications to the N-terminal tails of **histones** and the deposition of specific histone variants, have been implicated in BGC regulation. These epigenetic modifications influence chromatin density and can control tissue-specific and stress-responsive defence gene expression [59], making them plausible regulators of the tightly controlled spatiotemporal expression of BGCs.

Yu *et al.* reported that the thalianol and marneral BGCs in arabidopsis have distinct epigenetic architectures, characterised by an above-average association with trimethylation at lysine27 of histone H3 (H3K27me3), which marks **heterochromatic** regions, and the histone variant H2A.Z, which is commonly more associated with **euchromatic** regions (Figure 3A) [60]. Mutants compromised in H3K27me3 and H2A.Z deposition showed increased and decreased BGC transcription, respectively, without impacting the expression of nonclustered biosynthetic genes. Although the role of H2A.Z in gene expression continues to be debated [61,62], environmentally responsive genes (ERGs) are typically enriched with this H2 variant, suggesting that it has a role in their tissue-specific and/or stress-responsive expression [63]. Moreover, mutants in H2A.Z deposition (*arp6* and *hta9/hta11*) show reduced expression of both the thalianol and



Trends in Plant Science

Figure 3. Epigenetic regulation of biosynthetic gene cluster (BGC) expression. (A) Tri-methylation of H3K27 (H3K27me3) and deposition of histone variant H2A.Z leads to heterochromatic and euchromatic control over cluster expression, respectively. (B) In arabidopsis (*Arabidopsis thaliana)*, differential chromatin positioning between metabolite-producing and nonproducing tissues is associated with BGC expression. (C) Proposed model (adapted from [69]) illustrating the influence of the thalianol super-enhancer (SE) element on chromatin structure. The SE brings BGC promoters in proximity with transcription factors (TFs) and transcriptional co-activators to coordinate co-expression (i), which is prevented by a T-DNA insertion in the SE sequence (ii). (D) Abscisic acid (ABA)-induced expression of the *MARS* long noncoding (lnc)RNA decoys repressive LHP1 in PRC1 away from the marneral BGC, which reduces H3K27me3 formation and induces a chromatin loop that brings an ABA-responsive enhancer element (blue) in proximity with the BGC promoters to facilitate transcriptional induction; adapted from [72]. Figure created with BioRender (biorender.com).

marneral BGCs in roots, linked to an increased nucleosome occupancy in their coding and noncoding regions [64]. Together, these results suggest that the antagonising effects of H3K27me3 and H2A.Z on gene expression enable tightly coordinated spatiotemporal expression of triterpene BGCs.

Beyond arabidopsis, chromatin-dependent regulation of BGCs has been reported for both *Avina strigosa* (bristle oat) and *Zea mays* (maize). A comparison of the epigenetic makeup of the thalianol and marneral BGC in arabidopsis with the avenacin and benzoxazinoid BGCs in oat and maize, respectively, revealed that H3K27me3 enrichment is a conserved trait for BGCs across taxonomically unrelated plant species [60]. For the benzoxazinoid BGC, H3K27me3 enrichment correlated negatively with *Bx1*, *Bx3*, *Bx4*, and *Bx5* expression, suggesting that the age-dependent expression of this BGC is epigenetically controlled. Interestingly, aminophenoxazinone breakdown products of root-exudated benzoxazinoids inhibit histone deacetylases, promoting heterochromatin formation and stunting plant growth [65]. These epigenetic responses to benzoxazinoids not only contribute to their allelopathic activities, but may also explain the progressive silencing of the benzoxazinoid BGC with increasing plant age. In *A. strigosa*, cell-specific expression of the avenacin BGC correlates with chromatin decondensation in root epidermal cells, whereas the repressed state of this BGC in subepidermal and cortex cells is associated with a 28–65-fold compaction of the chromatin [66]. Together, these observations suggest that the highly spatiotemporal expression of BGCs in plants is at least partially *cis*-regulated by chromatin-based epigenetic processes.

Trans-regulation by topologically associated domains and lncRNAs

Beyond *cis*-regulatory histone modifications and variants in the transcriptional control of BGCs, emerging evidence from chromosomal conformation capture studies suggests a role for the 3D **chromatin topology**. The formation of **topologically associated domains (TADs)** allows for interactions between distant genomic regions [67]. In arabidopsis, the transcriptionally active thalianol BGC in root tissues becomes part of a TAD within the nuclear periphery (Figure 3B), which facilitates physical interactions with domains in the chromosomal arms and telomeres [68]. By contrast, the H3K27me3-enriched thalianol BGC in shoot tissues is located in the periphery of the nucleus, and topologically associates with transcriptionally silenced heterochromatic **pericentromeric** regions. A **super-enhancer** (SE) region within the thalianol cluster (Figure 3C) enables chromatin loop formation in root tissues, allowing for simultaneous interactions of all five BGC promoters with TFs and transcriptional coactivators [69]. Hence, SEs can have a role in the highly synchronised and tissue-specific expression of BGCs by controlling 3D chromatin topology [70].

Long noncoding RNAs (lncRNAs) are emerging as key regulators of chromatin structure and topology [71]. Roulé *et al.* reported that the arabidopsis lncRNA *MARS* controls ABA-dependent activity of the marneral BGC [72]. ABA-induced *MARS* transcripts bind to LHP1, an interactor of the polycomb repressive protein complexes PRC1 and PRC2 [73]. This decoy response reduces H3K27me3 deposition, resulting in a transcription-facilitating chromatin loop that brings *MARNERAL SYNTHASE 1* (*MRN1*) in proximity to an ABA-responsive enhancer element to induce BGC transcription (Figure 3D). Since lncRNA sequences are found in 73% of known plant BGCs, and lncRNA transcription correlates with stress-responsive BGC expression [72], it is plausible that lncRNAs have a conserved function in stress-responsive expression of BGCs.

To summarise, the highly coordinated, stress-responsive, and cell/tissue-specific expression of BGCs depends on not only phytohormones and TFs, but also *cis*- and *trans*-acting epigenetic mechanisms. Since stress-induced induction of plant immunity leads to changes in chromatin density and structure [74], these epigenetic responses allow for cost-efficient spatiotemporal activity of BGC-dependent defence. Considering the costs of inducible defences [75] and the ongoing evolutionary arms race between plants and their attackers [76], it is even possible that stress-responsive epigenetic processes have a role in the evolutionary formation and preservation of functionally related biosynthetic genes in BGCs.

Cluster evolution: from formation to selection

BGC evolution has been subject to ample discussion and speculation. While it is commonly assumed that gene duplication and neofunctionalisation drive BGC assembly [77], insight into the potential conditions and mechanisms driving gene duplication, as well as the selective forces responsible for consolidating and preserving BGCs, has only recently begun to emerge.

Transposable elements and histone signatures: evaluating a possible role in cluster evolution

TEs, broadly classified into 'class I' RNA-dependent 'retrotransposons' and 'class II' DNA transposons, are mobile elements that influence genome evolution [78]. It was recently shown that genomic integration of mobilised TEs is guided by the histone variant H2A.Z, which is enriched in environmentally responsive genes [79], including BGCs [60,64]. BGCs also typically reside in regions that are enriched with TEs [80–87] (Table 1), suggesting they have a role in BGC evolution. Both the avenacin BGC in oat and the dihydrolupeol BGC in *Lotus japonica* are enriched with TEs [84,85]. Moreover, a systematic study of BGC-associated terpenoid synthase (TS) and P450-dependent monooxygenases (CYP) across Eudicotidae revealed increased abundance of miniature inverted repeats transposons (MITEs) within the 100-kb flanking regions of these genes [87]. Interestingly, sequence analysis of MITEs associated with clustered and nonclustered TS and CYPs revealed that MITEs were recruited before cluster formation, indicating a structural role in cluster formation [87]. Synteny reconstruction by collinearity scanning has also implicated transposition and gene duplication in the segmental expansion and sequential insertion of ACS acyltransferases around a founding BAHD acyltransferase within the tomato acyl-sugar BGC [46]. Thus, the enrichment of BGCs with TEs and H2A.Z, combined with the highly improbable likelihood that functionally related genes co-cluster spontaneously, suggests an important role for TEs in the initial evolutionary steps toward cluster formation.

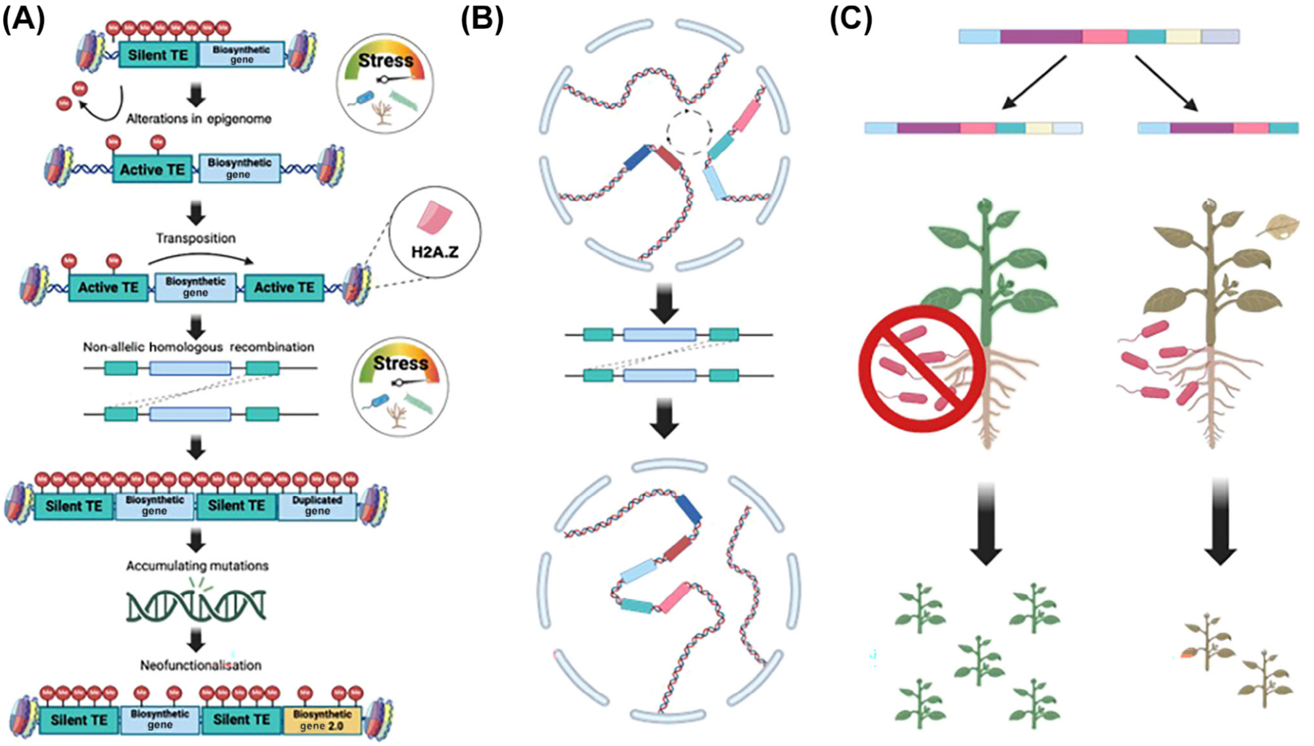
Under stress-free conditions, epigenetic silencing by DNA methylation, such as RNA-directed DNA- methylation (RdDM [88]), is crucial for suppressing TE activity. At the same time, plants actively antagonise DNA methylation in TE-rich regions by DNA demethylases, such as REPRESSOR

Table 1. Compendium of TE sequences associated with plant BGCs[a]

Species	Product	Class	Cluster length (kbp)	No. of genes	Location of TEs relative to BGC	TE class	No. of TEs within BGC	TE: gene ratio	Refs
Arabidopsis (*Arabidopsis thaliana*)	Thalianol	Triterpene	56	5	Within and peripheral	Class I and II	18	4.5	[80]
	Marneral	Triterpene	50	2	Within and peripheral	Class II			[80]
Oryza sativa	Hydroxycinnamoyl tyramine	Phenolamide	196	4	Within and peripheral		6	1.5	[81]
	Momilactone A	Diterpene	168	4	Within	Class I			[82,83]
Avina strigosa	Avenacin A-1	Triterpene	~245	12	Within and peripheral	Class I and II			[84]
Lotus japonicus	20-Hydroxy-betulinic acid	Triterpene	~103	5	Within		4	0.8	[85]
Papaver somniferum	Noscapine	Alkaloid	221	10	Within	Class I and II	23	2.3	[86]

[a] Blank cells indicate information that was not discernible from published literature.

OF SILENCING 1 (ROS1), which serves to limit proliferation of heterochromatin into nearby coding genes [89–91]. Independent studies have shown that biotic stress induces transient DNA demethylation, resulting in reduced TE silencing [74,92]. Accordingly, it is assumed that sustained exposure to stress over multiple generations increases the chance of TE mobilisation [93], causing structural and functional changes in the genome [94–96]. Since the histone variant H2A.Z guides TE insertion at ERGs [79], Wilkinson *et al.* [97] proposed that stress-induced mobilisation and integration of TEs facilitates the evolution and clustering of nucleotide-binding leucine-rich repeater proteins (NLRs), controlling effector-triggered immunity (ETI) against biotrophic pathogens [98,99]. Given that NLRs are under constant diversifying selection by rapidly coevolving pathogens [76,100], the effects of mobilised TEs on rapid NLR diversification explains how plants keep pace in the coevolutionary arms race with microbial pathogens. Akin to NLR clusters [101], BGCs are enriched with H2A.Z and TEs [60,64] (Table 1), suggesting that similar stress-induced transposition mechanisms contribute to BGC formation [102]. Alongside the mechanisms proposed by Wilkinson *et al.* [97], Smit and Lichman [102] recently proposed a model whereby topological associations of co-regulated genes precede BGC formation, increasing the likelihood of genomic



Trends in Plant Science

Figure 4. Refined model of biosynthetic gene cluster (BGC) evolution. (A) By integrating and expanding on the models proposed by Wilkinson *et al.* [97] and Smit and Lichman [102], we postulate that extended periods of stress repress epigenetic silencing of transposable elements (TEs), which facilitates their transposition and integration at H2A.Z-enriched biosynthetic genes. This stress-induced transposition, together with nonallelic homologous recombination between repetitive elements within TEs, leads to the duplication of biosynthetic genes. Cytosine methylation directed to regions containing newly inserted TEs enhances the C->T mutation rates, accelerating the divergence of redundant duplicates and facilitating neofunctionalisation and/or subfunctionalisation to expand chemodiversity. (B) As outlined by Smit and Lichman [102], functionally related genes within dynamic chromosomal regions become spatially associated within the 3D chromatin configuration of the genome. Subsequent ectopic recombination, transposition, and other genomic rearrangements lead to colocalisation of associated biosynthetic genes into the linear topology of BGCs, which we propose is catalysed by stress-inducible euchromatisation. (C) Retention of clusters as single units provides novel chemistry, which aids adaptive phenotypes, facilitates tight co-regulation, and promotes the inheritance of co-adaptive loci. Conversely, fragmentation prevents finely tuned co-regulation at the chromatin level, negating the adaptive value of such clusters in defence-related functions and increasing the risk of toxic build-up of intermediary products. The adaptive benefits in co-expression and ecological interactions, combined with the deleterious consequences of cluster fragmentation, act as selective forces to maintain linkage disequilibrium and ensure preservation of BGCs as single heritable units. Figure created with BioRender (biorender.com).

rearrangements that consolidate co-adaptive genes into linear clusters. In Figure 4 we integrate and develop these two emerging models, emphasising the catalytic role of stress-induced epigenetic changes in BGC formation, indicating how the desilencing of TEs and their integration into H2A.Z-enriched ERGs is enhanced following prolonged periods of stress to promote gene duplication, chemical diversification, and the clustering of functionally related genes.

Transposable elements: catalysts of adaptive chemical defences

The mobilisation of TEs and **retrotransposition** can facilitate gene duplication by **transduplication** and **retropositioning** [78,79,103]. The presence of repetitive sequences in TEs facilitates **nonallelic homologous recombination (NAHR)** [97], causing duplications, **exon shuffling**, and/or **gene fusions**, which give rise to new tandem arrays and gene variants (Figure 4A) [103]. Stress exposure increases NAHR in somatic plant cells [104,105], which can even be maintained into the next generation [106]. Somatic mutations by stress-induced TE mobilisation and NAHR can be transmitted into the gametes, particularly in plant species with late-sequestered and transient germlines [107]. Functional redundancy will then alleviate the selective constraints on newly formed gene duplicates, allowing for divergence and functionalisation [108]. This process will likely be accelerated by enhanced DNA methylation around newly inserted TE/biosynthetic gene conglomerates [109], since methylated cytosines are prone to deamination that result in G:C to A:T mutations [110]. Together, this interplay between dynamic stress-inducible epigenetic changes and genetic mutations advances neofunctionalisation and/or **subfunctionalisation**, leading to rapid chemical diversification and radiation of specialised metabolites with adaptive value.

In addition to TE mobilisation and NAHR, it is conceivable that the condensation of biosynthetic pathways into BGCs is preceded by nonlinear TAD-like higher-order chromatin interactions, facilitated by the presence of TEs, associated chromatin marks, and shared transcriptional machinery [102,111,112]. These 3D structures can form linear BGC topologies through **retropositioning,** and other genomic rearrangements, thereby bringing co-adaptive loci into physical proximity with each other [102,113,114]. Moreover, as proposed for rearrangements of NLRs [114,115], this process is likely to be accelerated by stress-induced hypomethylation and **ectopic recombination** with homologous TE sequences and H2A.Z-enriched gene promoters acting as TE integration and recombination hotspots [116] (Figure 4B). Reminiscent of the establishment of BGCs in Brassicaceae [117] and tandem duplication preceding BGC formation in *Nepeta* spp [118], our updated model predicts that stress-inducible epigenetic and genetic processes act in conjunction to drive metabolic evolution, diversifying enzymatic activity before BGC formation, with stress-induced genomic rearrangements consolidating linked enzymes into linear clusters.

Selection for BGCs

Yeaman [119] discusses how adaptive genetic architectures with clustered alleles of small effect may emerge in (sub)populations under strong selective pressures when migration and drift are low. However, genomic rearrangements that bring locally adaptive loci together have a more significant role in creating linkage disequilibrium, promoting the inheritance of co-adaptive loci in subsequent generations [113]. Furthermore, in addition to positive selection of BGCs with adaptive values in plant–biotic interactions, **genetic linkage** may be reinforced post clustering due to negative selection against cluster fragmentation, resulting in the accumulation of toxic intermediary products [120]. For example, the galactotransferase GAME1 in the tomatine BGC is essential for the conversion of the toxic precursors tomatidenol and tomatidine, thereby preventing self-toxicity [121]. In the avenacin BGC of oat, mutations in the late-acting genes *SAD3* and *SAD4* lead to root hair deficiency and membrane trafficking defects in root epidermal cells [122], whereas mutations in *UGT91G16* cause stunted root phenotypes [123] and compromised resistance to fungal root pathogens [124]. Interestingly, the positioning of genes within this BGC is in order with the stepwise biosynthesis of the pathway, which could explain why late-acting genes reside away from telomere to avoid deletions that cause toxic intermediates [84].

Thus, stress-induced euchromatisation may promote genomic rearrangements to expedite adaptation, promoting linkage of adaptive loci, while purifying selective pressures maintain genetic linkage to limit pathway fragmentation through disruptive recombination. In this regard, BGC evolution is a consequence of their unique function at the interface between endogenous epigenetic responses to stress and external adaptive ecological interactions, enabling tightly coordinated and cost-efficient control over powerful bioactive molecules driving these interactions.

Concluding remarks and future perspectives

Recent thinking has suggested that BGC evolution involves epigenetically controlled TE transposition and ectopic recombination, followed by external and endogenous selective pressures to preserve BGCs. The enrichment of BGCs with TEs and their positioning in dynamic chromosomal regions has likely facilitated the diversification of biosynthetic enzymes and radiation of adaptive metabolites. Consolidation of favourable gene combinations allows for co-regulation at the level of chromatin, ensuring tightly controlled, tissue-specific. and/or stress-responsive activity. BGC configurations then become fixed due to their adaptive value in plant–biotic interactions and maladaptive consequences of cluster fragmentation.

Genomic studies comparing functionally confirmed BGCs against pseudo-clusters could generate new insights into the genomic features facilitating the co-expression and evolution of BGCs. The discovery that specific TE families associate with BGCs would also justify research that aims to validate the model presented in Figure 4. For instance, the use of mutants in TE silencing and histone modifications/variant deposition in long-term experiments imposing environmental stresses across more than ten generations could reveal how a succession of stress-inducible TE activities, chromatin changes, and genome rearrangements shape the evolution of BGCs, their metabolites, and ecological activities. Such research requires a multidisciplinary approach, involving plant pathology, chemical ecology, molecular-(epi)genetics, and evolution biology, and would bestow great translational value, considering BGCs can generate chemistry of high economic and therapeutic value [125]. Moreover, this research can deliver knowledge to engineer a new generation of crop varieties that are better equipped to cope with the climatic extremes associated with climate change [126,127] (see also Outstanding questions).

Outstanding questions

Can the fitness benefits of metabolic gene clustering be characterised and defined in an ecologically relevant manner?

Does stress-related desilencing of TEs alter BGC-controlled secondary metabolism and drive defence-related interactions with beneficial insects and soil microbes?

Can the same BGCs control different metabolic products with distinct ecological roles, depending on tissue type, development, and environmental conditions?

Can heterologous expression systems help determine the functional benefits of clustering compared with nonclustered pathways?

To what extent do stress-related epigenetic activities, including changes in histone variants, noncoding RNAs, DNA methylation, and high-level chromatin organisation, control stress-responsive expression and accelerate the evolution of BGCs?

Author contributions

G.L.C. and J.T. wrote the manuscript.

Acknowledgements

We would like to thank the two anonymous reviewers for taking the time and effort to review an earlier draft of this manuscript. We are sincerely grateful for all the helpful comments and suggestions, which improved the quality of the final review. This work was supported by a BBSRC-IPA grant to J.T. (BB/W015250/1) and a supplementary grant from ENZA Zaden to J.T.

Declaration of interests

The authors declare that there are no competing interests associated with the manuscript.

References

1. Fang, C. *et al.* (2019) Exploring the diversity of plant metabolism. *Trends Plant Sci.* 24, 83–98
2. Kutchan, T.M. and Wurtzel, E.T. (2016) Plant metabolism, the diverse chemistry set of the future. *Science* 353, 1232–1236

3. Moore, B.M. *et al.* (2019) Robust predictions of specialized metabolism genes through machine learning. *Proc. Natl. Acad. Sci. U. S. A.* 116, 2344–2353
4. Weng, J.K. *et al.* (2012) The rise of chemodiversity in plants. *Science* 336, 1667–1670
5. Wink, M. (2003) Evolution of secondary metabolites from an ecological and molecular phylogenetic perspective. *Phytochemistry* 64, 3–19
6. Chan, E.K.F. *et al.* (2010) The complex genetic architecture of the metabolome. *PLoS Genet.* 6, e1001198
7. Waadt, R. (2020) Phytohormone signaling mechanisms and genetic methods for their modulation and detection. *Curr. Opin. Plant Biol.* 57, 31–40
8. Knaggs, A.R. (2003) The biosynthesis of shikimate metabolites. *Nat. Prod. Rep.* 20, 119–136
9. Weng, J.K. (2014) The evolutionary paths towards complexity: a metabolic perspective. *New Phytol.* 201, 1141–1149
10. Osbourn, A.E. (1996) Preformed antimicrobial compounds and plant defence against fungal attack. *Plant Cell* 8, 1821–1831
11. Lacchini, E. and Goossens, A. (2020) Combinatorial control of plant specialized metabolism: mechanisms, functions, and consequences. *Annu. Rev. Cell Dev. Biol.* 36, 291–313
12. Ton, J. *et al.* (2007) Priming by airborne signals boosts direct and indirect resistance in maize. *Plant J.* 49, 16–26
13. Mithöfer, A. and Boland, W. (2012) Plant defence against herbivores: chemical aspects. *Annu. Rev. Plant Biol.* 63, 431–450
14. Pierik, R. *et al.* (2014) Ecology of plant volatiles: taking a plant community perspective. *Plant Cell Environ.* 37, 1845–1853
15. Heil, M. (2008) Indirect defence via tritrophic interactions. *New Phytol.* 178, 41–61
16. Erb, M. *et al.* (2021) Using plant chemistry to improve interactions between plants, herbivores and their natural enemies: challenges and opportunities. *Curr. Opin. Biotechnol.* 70, 262–265
17. Piasecka, A. *et al.* (2015) Secondary metabolites in plant innate immunity: conserved function of divergent chemicals. *New Phytol.* 206, 948–964
18. Erb, M. and Kliebenstein, D.J. (2020) Plant secondary metabolites as defenses, regulators, and primary metabolites: the blurred functional trichotomy. *Plant Physiol.* 184, 39–52
19. Frey, M. *et al.* (1997) Analysis of a chemical plant defense mechanism in grasses. *Science* 277, 696–699
20. Foflonker, F. and Blaby-Haas, C.E. (2021) Colocality to cofunctionality: eukaryotic gene neighborhoods as a resource for function discovery. 38, 650–662
21. Kautsar, S.A. *et al.* (2017) PlantiSMASH: automated identification, annotation and expression analysis of plant biosynthetic gene clusters. *Nucleic Acids Res.* 45, 55–63
22. Huang, A. *et al.* (2019) A specialized metabolic network selectively modulates Arabidopsis root microbiota. *Science* 364, 546
23. Field, B. and Osbourn, A.E. (2008) Metabolic diversification–independent assembly of operon-like gene clusters in different plants. *Science* 320, 543–547
24. Liu, Z. *et al.* (2020) Formation and diversification of a paradigm biosynthetic gene cluster in plants. *Nat. Commun.* 11, 5354
25. Berendsen, R. *et al.* (2012) The rhizosphere microbiome and plant health. *Trends Plant Sci.* 17, 478–486
26. Sohrabi, R. *et al.* (2015) In planta variation of volatile biosynthesis. *Plant Cell* 27, 874–890
27. Frey, M. *et al.* (2009) Benzoxazinoid biosynthesis, a model for evolution of secondary metabolic pathways in plants. *Phytochemistry* 70, 1645–1651
28. Ahmad, S. *et al.* (2011) Benzoxazinoid metabolites regulate innate immunity against aphids and fungi in maize. *Plant Physiol.* 157, 317–327
29. Meihls, L. *et al.* (2013) Natural variation in maize aphid resistance is associated with 2,4-dihydroxy-7-methoxy-1,4-benzoxazin-3-one glucoside methyltransferase activity. *Plant Cell* 25, 2341–2355
30. Neal, A.L. *et al.* (2012) Benzoxazinoids in root exudates of maize attract *Pseudomonas putida* to the rhizosphere. *PLoS ONE* 7, e35498
31. Neal, A. and Ton, J. (2013) Systemic defense priming by *Pseudomonas putida* KT2440 in maize depends on benzoxazinoid exudation from the roots. *Plant Signal. Behav.* 8, 120–124
32. Cotton, T.E.A. *et al.* (2019) Metabolic regulation of the maize rhizobiome by benzoxazinoids. *ISME J.* 13, 1647–1658
33. Kudjordjie, E.N. *et al.* (2019) Maize synthesized benzoxazinoids affect the host associated microbiome. *Microbiome* 7, 59
34. Cadot, S. *et al.* (2021) Specific and conserved patterns of microbiota-structuring by maize benzoxazinoids in the field. *Microbiome* 9, 103
35. Hu, L. *et al.* (2018) Root exudate metabolites drive plant-soil feedbacks on growth and defense by shaping the rhizosphere microbiota. *Nat. Commun.* 9, 2738
36. Hu, L. *et al.* (2018) Plant iron acquisition strategy exploited by an insect herbivore. *Science* 361, 694–697
37. Itkin, M. *et al.* (2013) Biosynthesis of antinutritional alkaloids in solanaceous crops is mediated by clustered genes. *Science* 341, 175–179
38. Nützmann, H. *et al.* (2016) Plant metabolic clusters – from genetics to genomics. *New Phytol.* 211, 771–789
39. Nakayasu, M. *et al.* (2021) Tomato roots secrete tomatine to modulate the bacterial assemblage of the rhizosphere. *Plant Physiol.* 186, 270–284
40. Osbourn, A. *et al.* (2011) The saponins-polar isoprenoids with important and diverse biological activities. *Nat. Prod. Rep.* 28, 1261–1268
41. Morrissey, J. and Osbourn, A. (1999) Fungal resistance to plant antibiotics as a mechanism of pathogenesis. *Microbiol. Mol. Biol. Rev.* 63, 708–724
42. Quidde, T. *et al.* (1998) Detoxification of alpha-tomatine by *Botrytis cinerea*. *Physiol. Mol. Plant Pathol.* 52, 151–165
43. Ito, S. *et al.* (2004) Tomatidine and lycotetraose, hydrolysis products of α-tomatine by *Fusarium oxysporum* tomatinase, suppress induced defense responses in tomato cells. *FEBS Lett.* 571, 31–34
44. Fan, P. *et al.* (2020) Evolution of a plant gene cluster in Solanaceae and emergence of metabolic diversity. *eLife* 9, 1–26
45. Korenblum, E. *et al.* (2020) Rhizosphere microbiome mediates systemic root metabolite exudation by root-to-root signaling. *Proc. Natl. Acad. Sci. U. S. A.* 117, 3874–3883
46. Kerwin, R.E. *et al.* (2024) Tomato root specialized metabolites evolved through gene duplication and regulatory divergence within a biosynthetic gene cluster. *Sci. Adv.* 10, eadn3991
47. Kettle, A.J. *et al.* (2015) Degradation of the benzoxazolinone class of phytoalexins is important for virulence of *Fusarium pseudograminearum* towards wheat. *Mol. Plant Pathol.* 16, 946–962
48. Neilson, E.H. *et al.* (2013) Plant chemical defense: at what cost? *Trends Plant Sci.* 18, 250–258
49. Karban, R. *et al.* (1997) The benefits of induced defenses against herbivores. *Ecology* 78, 1351–1355
50. Méteignier, L. *et al.* (2023) Emerging mechanistic insights into the regulation of specialized metabolism in plants. *Nat. Plants* 9, 22–30
51. Hurst, L. *et al.* (2004) The evolutionary dynamics of eukaryotic gene order. *Nat. Rev. Genet.* 5, 299–310
52. Pieterse, C. *et al.* (2012) Hormonal modulation of plant immunity. *Annu. Rev. Cell Dev. Biol.* 28, 489–521
53. Wasternack, C. and Hause, B. (2013) Jasmonates: biosynthesis, perception, signal transduction and action in plant stress response, growth and development: an update to the 2007 review in Annals of Botany. *Ann. Bot.* 111, 1021–1058
54. Bai, Y. *et al.* (2021) Modulation of *Arabidopsis* root growth by specialized triterpenes. *New Phytol.* 230, 228–243
55. Pauwels, L. *et al.* (2010) NINJA connects the co-repressor TOPLESS to jasmonate signalling. *Nature* 464, 788–791
56. Nguyen, T.H. *et al.* (2023) A redundant transcription factor network steers spatiotemporal *Arabidopsis* triterpene synthesis. *Nat. Plants* 9, 926–937
57. Swinnen, G. *et al.* (2022) The basic helix–loop–helix transcription factors MYC1 and MYC2 have a dual role in the regulation of constitutive and stress-inducible specialized metabolism in tomato. *New Phytol.* 236, 911–928
58. Wisecaver, J.H. *et al.* (2017) A global coexpression network approach for connecting genes to specialized metabolic pathways in plants. *Plant Cell* 29, 944–959

59. Lämke, J. and Bäurle, I. (2017) Epigenetic and chromatin-based mechanisms in environmental stress adaptation and stress memory in plants. *Genome Biol.* 18, 124
60. Yu, N. *et al.* (2016) Delineation of metabolic gene clusters in plant genomes by chromatin signatures. *Nucleic Acids Res.* 44, 2255–2265
61. Zahraeifard, S. *et al.* (2018) Rice H2A.Z negatively regulates genes responsive to nutrient starvation but promotes expression of key housekeeping genes. *J. Exp. Bot.* 69, 4907–4919
62. Lei, B. and Berger, F. (2020) H2A variants in *Arabidopsis*: versatile regulators of genome activity. *Plant Commun.* 1, 100015
63. Coleman-Derr, D. and Zilberman, D. (2012) Deposition of histone variant H2A.Z within gene bodies regulates responsive genes. *PLoS Genet.* 8, e1002988
64. Nützmann, H. and Osbourn, A. (2015) Regulation of metabolic gene clusters in *Arabidopsis thaliana*. *New Phytol.* 205, 503–510
65. Venturelli, S. *et al.* (2015) Plants release precursors of histone deacetylase inhibitors to suppress growth of competitors. *Plant Cell* 27, 3175–3189
66. Wegel, E. *et al.* (2009) Cell type-specific chromatin decondensation of a metabolic gene cluster in oats. *Plant Cell* 21, 3926–3936
67. Dixon, J. *et al.* (2016) Chromatin domains: the unit of chromosome organization. *Mol. Cell* 62, 668–680
68. Nützmann, H. *et al.* (2020) Active and repressed biosynthetic gene clusters have spatially distinct chromosome states. *Proc. Natl. Acad. Sci. U. S. A.* 117, 13800–13809
69. Zhao, H. *et al.* (2022) Identification and functional validation of super-enhancers in *Arabidopsis thaliana*. *Proc. Natl. Acad. Sci. U. S. A.* 119, e2215328119
70. Shlyueva, D. *et al.* (2014) Transcriptional enhancers: from properties to genome-wide predictions'. *Nat. Rev. Genet.* 15, 272–286
71. Storz, G. (2002) An expanding universe of noncoding RNAs. *Science* 296, 1260–1263
72. Roulé, T. *et al.* (2022) The lncRNA MARS modulates the epigenetic reprogramming of the marneral cluster in response to ABA. *Mol. Plant* 15, 840–856
73. Veluchamy, A. *et al.* (2016) LHP1 regulates H3K27me3 spreading and shapes the three-dimensional conformation of the *Arabidopsis* genome. *PLoS ONE* 11, e0158936
74. Parker, A. *et al.* (2022) Epigenetics: a catalyst of plant immunity against pathogens. *New Phytol.* 233, 66–83
75. Heil, M. (2002) Ecological costs of induced resistance. *Curr. Opin. Plant Biol.* 5, 345–350
76. Jones, J. and Dangl, J. (2006) The plant immune system. *Nature* 444, 323–329
77. Polturak, G. *et al.* (2022) New and emerging concepts in the evolution and function of plant biosynthetic gene clusters. *Curr. Opin. Green Sustain. Chem.* 33, 100568
78. Seidl, M.F. and Thomma, B.P.H.J. (2017) Transposable elements direct the coevolution between plants and microbes. *Trends Genet.* 33, 842–851
79. Quadrana, L. *et al.* (2018) Transposon accumulation lines uncover histone H2A.Z-driven integration bias towards environmentally responsive genes. *bioRxiv*, Published online October 19, 2018. https://doi.org/10.1101/447870
80. Field, B. *et al.* (2011) Formation of plant metabolic gene clusters within dynamic chromosomal regions. *Proc. Natl. Acad. Sci. U. S. A.* 108, 16116–16121
81. Shen, S. *et al.* (2021) An Oryza-specific hydroxycinnamoyl tyramine gene cluster contributes to enhanced disease resistance. *Sci. Bull.* 66, 2369–2380
82. Wilderman, P. *et al.* (2004) Identification of syn-pimara-7,15-diene synthase reveals functional clustering of terpene synthases involved in rice phytoalexin/allelochemical biosynthesis. *Plant Physiol.* 135, 2098–2105
83. Shimura, K. *et al.* (2007) Identification of a biosynthetic gene cluster in rice for momilactones. *J. Biol. Chem.* 282, 34013–34018
84. Li, Y. *et al.* (2021) Subtelomeric assembly of a multi-gene pathway for antimicrobial defense compounds in cereals. *Nat. Commun.* 12, 2563–2513
85. Krokida, A. *et al.* (2013) A metabolic gene cluster in *Lotus japonicus* discloses novel enzyme functions and products in triterpene biosynthesis. *New Phytol.* 200, 675–690
86. Winzer, T. *et al.* (2012) A *Papaver somniferum* 10-gene cluster for synthesis of the anticancer alkaloid noscapine. *Science* 336, 1704–1708
87. Boutanaev, A.M. and Osbourn, A.E. (2018) Multigenome analysis implicates miniature inverted-repeat transposable elements (MITEs) in metabolic diversification in eudicots. *Proc. Natl. Acad. Sci. U. S. A.* 115, E6650–E6658
88. Zhang, H. *et al.* (2018) Dynamics and function of DNA methylation in plants. *Nat. Rev. Mol. Cell Biol.* 19, 489–506
89. Tang, K. *et al.* (2016) The DNA demethylase ROS1 targets genomic regions with distinct chromatin modifications. *Nat. Plants* 2, 16169
90. Du, X. *et al.* (2023) Molecular basis of the plant ROS1-mediated active DNA demethylation. *Nat. Plants* 9, 271–279
91. Dowen, R.H. *et al.* (2012) Widespread dynamic DNA methylation in response to biotic stress. *Proc. Natl. Acad. Sci. U. S. A.* 109, E2183–E2191
92. Wessler, S.R. (1996) Turned on by stress. Plant retrotransposons. *Curr. Biol.* 6, 959–961
93. Galindo-González, L. *et al.* (2017) LTR-retrotransposons in plants: engines of evolution. *Gene* 626, 14–25
94. Lisch, D. (2013) How important are transposons for plant evolution? *Nat. Rev. Genet.* 14, 49–61
95. Quadrana, L. *et al.* (2019) Transposition favors the generation of large effect mutations that may facilitate rapid adaption. *Nat. Commun.* 10, 3421–3410
96. Roquis, D. *et al.* (2021) Genomic impact of stress-induced transposable element mobility in Arabidopsis. *Nucleic Acids Res.* 49, 10431–10447
97. Wilkinson, S.W. *et al.* (2019) Surviving in a hostile world: plant strategies to resist pests and diseases. *Annu. Rev. Phytopathol.* 57, 505–529
98. Cui, H. *et al.* (2015) Effector-triggered immunity: from pathogen perception to robust defense. *Annu. Rev. Plant Biol.* 66, 487–511
99. Michelmore, R.W. and Meyers, B.C. (1998) Clusters of resistance genes in plants evolve by divergent selection and a birth-and-death process'. *Genome Res.* 8, 1113–1130
100. Chae, E. *et al.* (2014) Species-wide genetic incompatibility analysis identifies immune genes as hot spots of deleterious epistasis. *Cell* 159, 1341–1351
101. Van Wersch, S. and Li, X. (2019) Stronger when together: clustering of plant NLR disease resistance genes. *Trends Plant Sci.* 24, 688–699
102. Smit, S.J. and Lichman, B.R. (2022) Plant biosynthetic gene clusters in the context of metabolic evolution. *Nat. Prod. Rep.* 39, 1465–1482
103. Krasileva, K.V. (2019) The role of transposable elements and DNA damage repair mechanisms in gene duplications and gene fusions in plant genomes. *Curr. Opin. Plant Biol.* 48, 18–25
104. Hohn, B. *et al.* (2002) Pathogen stress increases somatic recombination frequency in *Arabidopsis*. *Nat. Genet.* 30, 311–314
105. Kovalchuk, I. *et al.* (2003) Pathogen-induced systemic plant signal triggers DNA rearrangements. *Nature* 423, 760–762
106. Molinier, J. *et al.* (2006) Transgeneration memory of stress in plants. *Nature* 442, 1046–1049
107. Reusch, T.B.H. *et al.* (2021) Evolution via somatic genetic variation in modular species. *Trends Ecol. Evol.* 36, 1083–1092
108. Panchy, N. *et al.* (2016) Evolution of gene duplication in plants. *Plant Physiol.* 171, 2294–2316
109. Wang, Z. and Baulcombe, D. (2020) Transposon age and non-CG methylation. *Nat. Commun.* 11, 122
110. Ossowski, S. *et al.* (2010) Rate and molecular spectrum of spontaneous mutations in *Arabidopsis thaliana*. *Science* 327, 92–94
111. Sun, L. *et al.* (2020) Heat stress-induced transposon activation correlates with 3D chromatin organization rearrangement in *Arabidopsis*. *Nat. Commun.* 11, 1886
112. Lanctôt, C. *et al.* (2007) Dynamic genome architecture in the nuclear space: regulation of gene expression in three dimensions. *Nat. Rev. Genet.* 8, 104–115
113. Yeaman, S. (2013) Genomic rearrangements and the evolution of clusters of locally adaptive loci. *Proc. Natl. Acad. Sci. U. S. A.* 110, E1743–E1751

114. Zhang, J. (2003) Evolution by gene duplication: an update. *Trends Ecol. Evol.* 18, 292–298
115. Boyko, A. *et al.* (2007) Transgenerational changes in the genome stability and methylation in pathogen-infected plants. *Nucleic Acids Res.* 35, 1714–1725
116. Choi, K. *et al.* (2013) *Arabidopsis* meiotic crossover hot spots overlap with H2A.Z nucleosomes at gene promoters. *Nat. Genet.* 45, 1327–1336
117. Liu, Z. *et al.* (2020) Drivers of metabolic diversification: how dynamic genomic neighbourhoods generate new biosynthetic pathways in the Brassicaceae. *New Phytol.* 227, 1109–1123
118. Lichman, B.R. *et al.* (2020) The evolutionary origins of the cat attractant nepetalactone in catnip. *Sci. Adv.* 6, eaba0721
119. Yeaman, S. (2022) Evolution of polygenic traits under global vs local adaptation. *Genetics* 220, iyab134
120. Gierl, A. and Frey, M. (2001) Evolution of benzoxazinone biosynthesis and indole production in maize. *Planta* 213, 493–498
121. Itkin, M. *et al.* (2011) GLYCOALKALOID METABOLISM1 is required for steroidal alkaloid glycosylation and prevention of phytotoxicity in tomato. *Plant Cell* 23, 4507–4525
122. Mylona, P. *et al.* (2008) Sad3 and Sad4 are required for saponin biosynthesis and root development in oat. *Plant Cell* 20, 201–212
123. Orme, A. *et al.* (2019) A noncanonical vacuolar sugar transferase required for biosynthesis of antimicrobial defense compounds in oat. *Proc. Natl. Acad. Sci. U. S. A.* 116, 27105–27114
124. Papadopoulou, K. *et al.* (1999) Compromised disease resistance in saponin-deficient plants. *Proc. Natl. Acad. Sci. U. S. A.* 96, 12923–12928
125. Stephenson, M.J. and Osbourn, A. (2020) Making drugs out of sunlight and ‘thin air’: an emerging synergy of synthetic biology and natural product chemistry. *Biochemist* 42, 34–39
126. Clemensen, A.K. *et al.* (2020) Ecological implications of plant secondary metabolites - phytochemical diversity can enhance agricultural sustainability. *Front. Sustain. Food Syst.* 4, 547826
127. Gfeller, V. *et al.* (2023) Plant secondary metabolite-dependent plant-soil feedbacks can improve crop yield in the field. *eLife* 12, e84988