Active and repressed biosynthetic gene clusters have spatially distinct chromosome states

Hans-Wilhelm Nützmann[a,1], Daniel Doerr[b], América Ramírez-Colmenero[c], Jesús Emiliano Sotelo-Fonseca[c], Eva Wegel[d], Marco Di Stefano[e], Steven W. Wingett[f], Peter Fraser[g], Laurence Hurst[a], Selene L. Fernandez-Valverde[c], and Anne Osbourn[h]

[a]Department of Biology and Biochemistry, The Milner Centre for Evolution, University of Bath, BA2 7AY Bath, United Kingdom; [b]Faculty of Technology and Center for Biotechnology, Bielefeld University, 33615 Bielefeld, Germany; [c]Unidad de Genómica Avanzada, Langebio, Centro de Investigación y Estudios Avanzados del Instituto Politécnico Nacional, Irapuato, Guanajuato 36824, Mexico; [d]Bioimaging, John Innes Centre, NR4 7UH Norwich, United Kingdom; [e]Centre Nacional d'Anàlisi Genòmica, Centre for Genomic Regulation (CNAG-CRG), The Barcelona Institute of Science and Technology (BIST), 08028 Barcelona, Spain; [f]Bioinformatics, Babraham Institute, CB22 3AT Cambridge, United Kingdom; [g]Department of Biological Science, Florida State University, Tallahassee, FL 32306; and [h]Department of Metabolic Biology, John Innes Centre, NR4 7UH Norwich, United Kingdom

Edited by Dominique C. Bergmann, Stanford University, Stanford, CA, and approved April 24, 2020 (received for review November 26, 2019)

While colocalization within a bacterial operon enables coexpression of the constituent genes, the mechanistic logic of clustering of nonhomologous monocistronic genes in eukaryotes is not immediately obvious. Biosynthetic gene clusters that encode pathways for specialized metabolites are an exception to the classical eukaryote rule of random gene location and provide paradigmatic exemplars with which to understand eukaryotic cluster dynamics and regulation. Here, using 3C, Hi-C, and Capture Hi-C (CHi-C) organ-specific chromosome conformation capture techniques along with high-resolution microscopy, we investigate how chromosome topology relates to transcriptional activity of clustered biosynthetic pathway genes in *Arabidopsis thaliana*. Our analyses reveal that biosynthetic gene clusters are embedded in local hot spots of 3D contacts that segregate cluster regions from the surrounding chromosome environment. The spatial conformation of these cluster-associated domains differs between transcriptionally active and silenced clusters. We further show that silenced clusters associate with heterochromatic chromosomal domains toward the periphery of the nucleus, while transcriptionally active clusters relocate away from the nuclear periphery. Examination of chromosome structure at unrelated clusters in maize, rice, and tomato indicates that integration of clustered pathway genes into distinct topological domains is a common feature in plant genomes. Our results shed light on the potential mechanisms that constrain coexpression within clusters of nonhomologous eukaryotic genes and suggest that gene clustering in the one-dimensional chromosome is accompanied by compartmentalization of the 3D chromosome.

gene cluster | Hi-C | *Arabidopsis thaliana* | H3K27me3 | chromosome organization

Gene order is a central feature that distinguishes eukaryotic genomes from their prokaryotic counterparts. Prokaryotic genomes are characterized by colocalization of functionally related genes in operons (1). In contrast, functionally related genes in eukaryotes are commonly dispersed throughout genomes. Progress in genomics and transcriptomics have shown that gene order in eukaryotes is far from random and that the positioning of genes affects their transcriptional activity and evolutionary retention (2–4). In addition, diverse examples of colocalized and functionally related genes ("operon-like" gene clusters) have been identified in eukaryotes that are reminiscent of gene organization in prokaryotes (5–11). While the polycistronic transcription of bacterial operons provides an immediate and established mechanistic logic for coexpression of functionally related genes, for eukaryotic operon-like clusters of genes that are predominantly transcribed as single monocistronic units with individual promoters, this mechanistic logic is not obvious (1, 3, 12).

In plants, it has recently been discovered that genomes contain regions characterized by operon-like clusters of colocalized and non-sequence-related genes involved in the biosynthesis of natural products. These clusters encode pathways for the biosynthesis of diverse molecules, ranging from medicinal alkaloids to polyketide components of wax layers and triterpenes that shape the root microbiota (13–16). The identification of this clustering phenomenon has led to the development of new genomics-driven strategies for pathway discovery (17–20). Plant biosynthetic gene clusters range in size from ∼35 kb to several hundred kb. They are located in genomic regions that are prone to chromosomal rearrangement and have arisen by recruitment of genes from elsewhere in the genome followed by neofunctionalization (14, 21–23). Individual metabolic gene clusters and their variants are usually confined to narrow taxonomic windows (14).

The genes within these biosynthetic clusters are typically coexpressed in specific plant organs and/or in response to certain environmental triggers. High transcriptional activity in metabolite producing cells is often contrasted by tight transcriptional silencing in nonproducing cells (14, 19, 24, 25).

Significance

Clusters of coexpressed and colocalized biosynthetic pathway genes in plants are a paradigmatic example of the nonrandom organization of the eukaryotic genome and present an ideal opportunity to understand the logic of eukaryote gene cluster regulation. Here, we carry out an in-depth analysis of the chromosomal topology of biosynthetic gene clusters and their positioning in nuclear space. We demonstrate that plant biosynthetic gene clusters reside in highly interactive domains that undergo marked changes in local conformation and nuclear positioning in cluster expressing and nonexpressing organs. As such, metabolic gene clusters rank among the most dynamic regions in the genome of the model species *A. thaliana*. Our results shed light on the potential mechanisms that constrain coexpression within clusters of genes.

Author contributions: H.-W.N. and A.O. designed research; H.-W.N., D.D., A.R.-C., J.E.S.-F., E.W., M.D.S., and S.L.F.-V. performed research; P.F. contributed new reagents/analytic tools; H.-W.N., D.D., A.R.-C., J.E.S.-F., M.D.S., S.W.W., and S.L.F.-V. analyzed data; and H.-W.N., L.H., S.L.F.-V., and A.O. wrote the paper.

The authors declare no competing interest.

This article is a PNAS Direct Submission.

Data deposition: The sequences reported in this paper have been deposited in the National Center for Biotechnology Information (NCBI) Gene Expression Omnibus (GEO) database, https://www.ncbi.nlm.nih.gov/ (accession no. PRJNA576277).

[1]To whom correspondence may be addressed. Email: h.nuetzmann@bath.ac.uk.

This article contains supporting information online at https://www.pnas.org/lookup/suppl/doi:10.1073/pnas.1920474117/-/DCSupplemental.

First published June 3, 2020.

It is hypothesized that physical linkage of functionally related genes in eukaryotes is associated with specialized regulatory processes (3, 12). Physical linkage may facilitate coordinate gene regulation through shared promoter and regulatory DNA elements as well as common epigenetic modifications of cluster-associated histones and DNA motifs (19, 20, 26–31). Furthermore, it has been proposed that 3D chromosome structure and localization to specific nuclear territories are mechanisms for coordinate transcriptional regulation of adjacent genes. Seminal work on clusters of homologous genes in humans and animals, such as the HOX and β-globin clusters, supports this hypothesis (32–34).

In plants and for eukaryotic clusters of nonhomologous genes, in general, however, it remains unknown how groups of neighboring and coexpressed genes integrate into the nuclear 3D environment. A recent study in the filamentous fungus *Epichloë festucae* reported the localization of a biosynthetic gene cluster in a single topologically associated domain (TAD), and the authors suggested that activation of this cluster may be associated with a remodeling of this chromosome structure (35).

In previous studies, we have shown that common signatures of chromatin marks delineate plant biosynthetic gene clusters and that *A. thaliana* chromatin mutants have altered cluster transcript levels compared to the wild type (20). Furthermore, in diploid oat (*Avena strigosa*), we have shown by high-resolution DNA in situ hybridization that expression of a biosynthetic gene cluster for the synthesis of antimicrobial defense compounds known as avenacins is associated with chromatin decondensation (25).

Here, we characterize the chromosome topology of metabolic gene clusters in plants using previously characterized gene clusters in *A. thaliana* as our models. We have recently shown that these clustered biosynthetic pathways form a metabolic network that shapes the root microbiota (13). These gene clusters are coordinately transcribed in the roots but silenced in the aerial organs of the plant (20, 23, 24, 36). They, therefore, offer an ideal experimental system for investigating the organ-specific regulation of plant biosynthetic gene clusters.

To define the chromosome architecture of metabolic gene clusters and their integration into the nuclear environment, we carried out chromosome conformation capture (3C, Hi-C, and CHi-C) experiments using nuclear preparations from roots and leaves of *A. thaliana* seedlings. We show that the *A. thaliana* biosynthetic gene clusters are embedded in local interactive 3D chromosomal domains that adopt different structures in expressing and nonexpressing organs. Comparative analysis reveals that these domains undergo some of the most drastic genome-wide changes to chromosome topology when roots and leaves are compared. We further demonstrate that these biosynthetic clusters are localized to heterochromatic areas of the genome when silenced. Incorporation and analysis of available Hi-C maps implicate histone H3 lysine 27 trimethylation (H3K27me3) as a central feature of the 3D domains at silenced clusters. Examination of chromosome structure at unrelated clusters in maize, rice, and tomato indicates that integration of clustered pathway genes into distinct topologically associated domains is a widespread feature in plants.

Collectively, our work provides a high-resolution view of the nuclear organization of biosynthetic gene clusters in plants. It demonstrates that a unique pattern of chromosomal conformations is established at clusters. Our findings also open up a potential route to manipulate plant specialized metabolism by interfering with higher-order regulatory mechanisms

Results and Discussion

Analysis of 3D Chromosome Conformation Reveals Organ-Specific Differences. To define the 3D chromosome architecture at biosynthetic gene clusters in *A. thaliana*, we set out to establish chromosome conformation capture protocols in conditions that would reflect transcriptional "on" and "off" states of the clusters. In earlier work, we and others have shown that expression of several previously characterized biosynthetic gene clusters, among them the thalianol cluster, is tightly repressed in the leaves and highly expressed in the roots of young *A. thaliana* seedlings (20, 23, 24, 36, 37). To corroborate this, we performed whole transcriptome analysis of RNA extracted from roots and leaves of 7-d-old seedlings (Dataset S1). We observed marked changes in the transcript levels for three distinct clusters—the thalianol, marneral, and arabidiol/baruol clusters—when root and leaves were compared (Fig. 1 *A* and *B*).

We then performed genome-wide Hi-C analysis for DNA from both organs. We obtained 71–108 million valid unique paired-end reads from each library (*SI Appendix*, Tables S1 and S2). We corrected and normalized the derived interaction counts for experimental biases and genomic distance. The normalized counts serve as a measure of interaction strength between any two chromosomal sites and were plotted as 2D Hi-C maps. Visual examination showed strong interchromosomal contacts between pericentromeric blocks and between all telomers in both roots and leaves (Fig. 1*C*). We could also readily detect a number of prominent off-diagonal punctate signals that reflect interactions of defined chromosomal loci known as interactive heterochromatic islands (IHIs) (Fig. 1*C* and *SI Appendix*, Fig. S1) (38, 39). These features are overall consistent with those observed in the previously reported overarching conformation of *A. thaliana* chromosomes in whole seedlings (38–40).

In in-depth comparative examinations of the chromosome architecture, we observed several differences between the chromosome features observed for roots and leaves. In leaves, significantly increased interaction counts were enriched in the pericentromeric regions while in roots enhanced interactions were predominantly localized to the chromosome arms and telomeres (Fig. 1 *C* and *D* and *SI Appendix*, Fig. S1 and Dataset S2). We further detected changes to the IHIs. The intensity of individual interchromosomal IHI interactions varied significantly, and one previously undescribed IHI located on chromosome 1 was identified (Fig. 1*D* and *SI Appendix*, Figs. S2 and S3). These findings highlight organ-specific reconfigurations of the *A. thaliana* 3D chromosome architecture. Differences in interaction intensity between different tissues have also been reported for Hi-C analyses in rice and maize (41).

In *A. thaliana*, alterations to interaction intensities of pericentromeric and telomeric regions of the genomes and IHIs have been described for mutant lines that are defective for different epigenetic pathways, such as DNA methylation and histone H3 lysine 9 methylation (38). This may suggest that differences in the epigenetic environment between roots and leaves underlie the observed 3D changes. Furthermore, differences in nuclear shape and ploidy levels of individual nuclei of roots and leaves may be associated with the variations in interaction intensities (42).

Silenced Metabolic Gene Clusters Associate with Heterochromatic Areas within the Nucleus. We then focused our analysis on the thalianol metabolic gene cluster. We have recently shown that the metabolic products derived from the thalianol cluster have important roles in shaping the root microbiome of *A. thaliana* (13). We have further established that the cluster is delineated by repressive H3K27me3 marks and the histone variant H2A.Z involved in positive regulation (20, 37). The thalianol cluster consists of four core genes that cover ∼33 kb and a peripheral gene that is separated from the core cluster by 10 kb (the latter 10-kb region including two unrelated intervening genes) (Fig. 1*A*) (13, 24). All cluster genes are widely expressed in root tissues and repressed in aerial plant tissues (Fig. 1*B* and *SI Appendix*, Fig. S4).

Visual inspection of our Hi-C maps for roots and leaves show small local interactive domains encompassing the thalianol cluster and separating it from the neighboring genomic environment (Fig. 2*A* and *SI Appendix*, Fig. S5). Reanalysis of previous whole

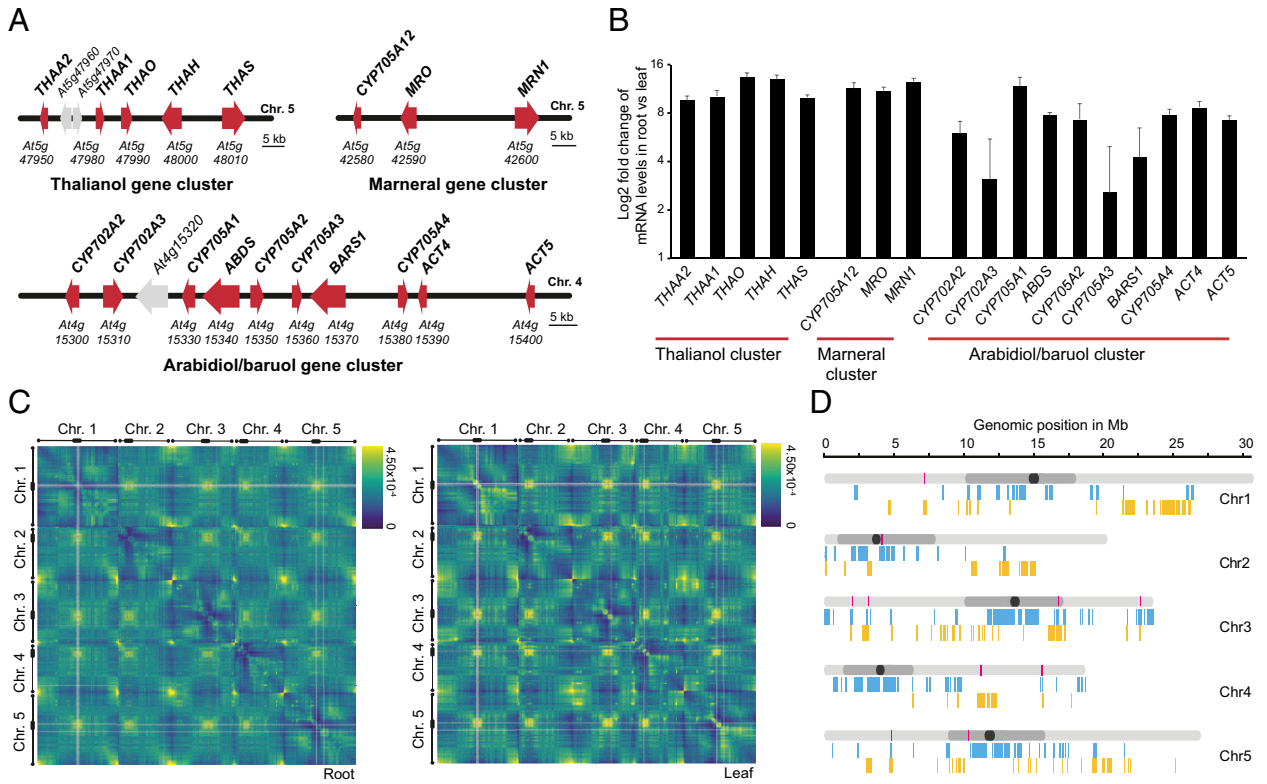

Fig. 1. Organ-specific gene cluster expression and chromosome conformation in *A. thaliana*. (*A*) The thalianol, marneral, and baruol/arabidiol gene clusters. Red arrows, cluster genes. Gray arrows, uncharacterized genes. (*B*) Relative quantification of mRNA levels of thalianol, marneral, and arabidiol/baruol cluster genes in the roots and leaves of 7-d-old seedlings as assessed by RNA-sequencing (RNA-seq) analysis of three biological replicates. Error bars indicate the SE of logFC. (*C*) 2D Hi-C interaction maps of the 3D conformation of *A. thaliana* chromosomes in roots and leaves. Chromosomes are labeled from *Left* to *Right* and *Top* to *Bottom*. Centromeric and pericentromeric regions are marked with black rounded boxes. Telomeres are marked with circles. Yellow to blue coloring indicates strong to weak interaction tendency. Genomic bin size: 25 kb. (*D*) Location of differentially interacting regions on *A.thaliana* chromosomes. Distribution of the top 10% regions engaging in most differential interactions throughout the genome are shown (P value < 0.01). Regions with increased interaction tendency in leaves and roots are shown in blue and yellow, respectively. Note: Regions may interact with different chromosomal sites in root and leaf and may, thus, show both blue and yellow markings. Light gray, chromosome arms; dark gray, pericentromeric region; black, centromeric region; pink, IHIs.

seedling-derived conformation data shows a similar domain at the thalianol cluster (*SI Appendix*, Fig. S6) (31). Strikingly, the location and interaction strength of the interactive domain change between roots and leaves, and two distinct interactive domains can be distinguished (Fig. 2*A*). An area encompassing the thalianol cluster engages in very strong 3D contacts in roots while, in leaves, the location of the interactive domain shifts downstream and covers the cluster and a region downstream of the cluster. This domain is larger than in roots, and the intensity of domain-wide 3D interactions is reduced (Fig. 2*A*). We then performed A/B compartment analyses for all chromosome arms of our Hi-C maps and analyzed the compartment association of the thalianol cluster. In leaves, we found the cluster to be positioned within a B compartment, a more compact structural domain with increased intradomain contacts, and, in roots, we found the cluster to be localized within an A compartment with depleted intradomain contacts and a looser structure (*SI Appendix*, Fig. S7).

After detecting these differences, we revisited the comparative analysis of our Hi-C libraries to identify differential interactions associated with the cluster. By calling genome-wide significant differential interactions, we observed a striking pattern of both local and global changes to the 3D structure of the thalianol cluster. Our comparative analysis shows that the chromosomal region encompassing the thalianol cluster ranks among the most differentially interacting areas of the *A. thaliana* genome in the root–leaf comparison (Fig. 2*B*). In leaves, the cluster engages in significantly enriched interactions toward the pericentromeric areas of the chromosome (Fig. 2*C* and *SI Appendix*, Fig. S8). In contrast, root-specific interactions are significantly elevated toward the long arm of chromosome 5 outside the pericentromeric region (Fig. 2*C* and Dataset S3).

We identified similar conformational features for the marneral and arabidiol/baruol gene clusters (Fig. 2*B* and *SI Appendix*, Fig. S9). These patterns reflect the different transcriptional states of the clusters. In leaves, when the biosynthetic gene clusters are silenced, the clusters are directed toward heterochromatic areas of the genome, and in roots, when active, are located toward open transcriptionally active areas of the genome. Accordingly, the genes located in regions differentially interacting with metabolic gene clusters show significantly lower expression levels in leaves compared to roots (Fig. 2*D*).

To corroborate the change in localization, we performed 3D DNA fluorescence in situ hybridization (FISH) analysis of the thalianol cluster. Chromocenters are preferentially associated with the nuclear periphery in *A. thaliana* nuclei (43, 44), and, as such, we analyzed cluster localization toward the periphery in nuclei of roots and leaves. We show that the cluster region strongly associates with the nuclear periphery in leaf tissues while, in roots, this association is significantly reduced (Fig. 2*E*).

Conformational Switching Accompanies Changes in Transcriptional Activity at the Thalianol Metabolic Gene Cluster. To better define the changes to the 3D chromosome architecture at metabolic gene clusters, we established a CHi-C protocol for *A. thaliana* (45, 46).

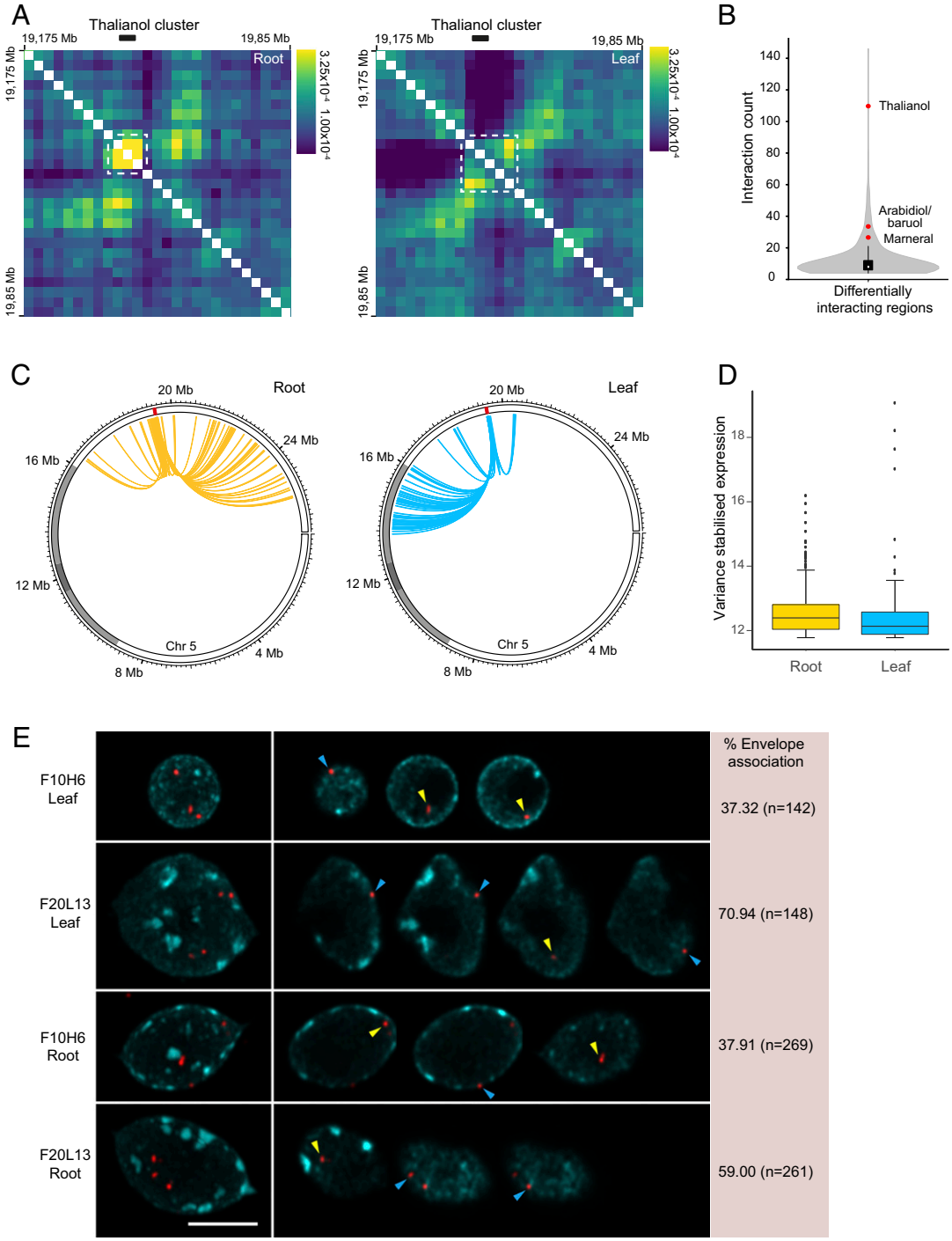

Fig. 2. Integration of the thalianol gene cluster in the 3D nuclear space. (*A*) Detailed 2D Hi-C maps of the genomic region surrounding the thalianol cluster. White boxes indicate cluster-associated local interactive domains in roots (*Left*) and leaves (*Right*). Chromosome coordinates are labeled from *Left* to *Right* and *Top* to *Bottom*. Yellow to blue coloring indicates strong to weak interaction tendency. Genomic bin size: 25 kb. (*B*) Violin plot of frequency and range of interaction counts of differentially interacting regions. The thalianol, arabidiol/baruol, and marneral cluster-associated interactive domains are indicated by the red dots. (*C*) Circos plots showing intrachromosomal differential genomic interactions for the thalianol cluster in leaves and roots. Chromosome arms are shown in white, pericentromers are shown in light gray, centromeres are shown in dark gray, and the thalianol cluster is shown in red. Each arc/connection represents a significantly enriched interaction ($P < 0.01$). (*D*) Box plots showing normalized expression levels of genes within chromosomal regions of differential interaction with the thalianol, marneral, and arabidiol/baruol clusters. Variance stabilized gene expression values are shown for differentially interacting regions in roots and leaves. Gene expression levels are significantly higher in roots compared to leaves ($P = 1.118\times10^{-7}$, Wilcoxon rank-sum test). (*E*) The 3D DNA FISH analysis of nuclear envelope association of the thalianol cluster. Representative images and envelope association percentage for bacterial artificial chromosomes BAC F20L13 (covering the thalianol cluster) and a control, BAC F10H6 (covering an area of open chromatin). The nuclei on the *Left* are maximum intensity projections of a *z* stack through each nucleus. On the *Right*, each locus is shown in a single optical section through the nucleus in the *xy* direction. Note, the higher envelope association rate of F20L13 compared to F10H6 in leaves ($P = 1.11\times10^{-8}$, two-sided Fisher's exact test) and the significantly lower association of F20L13 in roots compared to leaves ($P = 0.019$, two-sided Fisher's exact test). (Scale bar, 5 μm.) Blue arrow, foci associated with the nuclear envelope; yellow arrows, foci not associated with nuclear envelope. n, number of foci counted.

We defined a set of genomic regions with sizes between 200 and 600 kb and designed capture probes that cover all restriction fragments within. We chose the regions based on the annotation of metabolic gene clusters and distribution along all chromosomes and central chromosome features (Dataset S4). We obtained libraries with similar yields of valid di-tags but with much improved sequence depth at the captured sites compared to our Hi-C library and previously published data (*SI Appendix*, Tables S3–S6).

Analysis of the CHi-C library enabled us to recapitulate the chromosome-wide interaction switch of the thalianol cluster (*SI Appendix*, Fig. S10). Furthermore, it enabled us to precisely analyze the local variations in the 3D chromosome structure at the thalianol cluster coinciding with the transcriptional on and off states. In roots, when active, the cluster is located within an interactive domain that consists of two layers of variable strength. The smaller and higher intensity domain precisely demarcates the thalianol cluster, ranging from *THAA2*, the peripheral cluster gene, to the *THAS* gene with a size of 50 kb. The larger domain extends from the noncluster gene *At5g47910* to the thalianol cluster (Fig. 3*A* and *SI Appendix*, Fig. S11*A*).

In contrast to the root-specific domain, the interactive domain formed in leaves, when the cluster is silenced, is larger in size and starts at the *THAA2* gene, the peripheral cluster gene, and ends at the nonclustered and noncoexpressed genes *At5g48150* and *At5g48160*. Overall, the domain is 110 kb in size and covers the thalianol cluster and a group of genes with increased expression level in roots compared to leaves (*At5g48070* to *At5g48140*) (Fig. 3*B* and *SI Appendix*, Fig. S11 *A* and *B*).

Structural modeling of the major 3D domains associated with the thalianol cluster indicates that, when active, the locus assumes a compact conformation and, when silent, it is incorporated into a chromosomal loop (*SI Appendix*, Fig. S12).

Differential visualization of CHi-C maps of root and leaves shows a striking border between the silencing and the activating domains around the gene *At5g48050*, an area depleted of obvious regulatory elements (Fig. 3*C*). Independent 3C experiments that measure contact intensity between individual restriction fragments corroborate the identified domain structure (Fig. 3*E*).

This association with two local interactive domains, i.e., a bimodal chromosomal configuration, may allow the thalianol cluster to read regulatory information from two distinct chromosomal areas (Fig. 3*D*). For the HoxD cluster in mice, location between two TADs and dynamic association with either of them ensures collinear distribution of Hox transcription factors to the correct developmental body structures (47).

It is striking that the chromocenter interactions of the clusters are partly driven by contacts of the cluster downstream region that are specific for the silencing domain (*SI Appendix*, Fig. S10). We suggest a threefold mechanism in which the cluster is brought in proximity to the envelope: a) the cluster is released from an interactive domain associated with strong transcriptional activity; b) establishment of new contacts with a region downstream of the cluster; c) positioning toward heterochromatic chromosomal areas near the nuclear periphery.

Similar to the thalianol cluster, the 3D chromosome architecture at the marneral cluster is associated with a complex pattern of local and regional interactions that differ between roots and leaves. As seen in our Hi-C maps, a large regional interactive domain is formed between the marneral cluster and a genomic area 300 kb away from the cluster. This domain is specifically formed in leaf organs when the cluster is silenced (*SI Appendix*, Figs. S13 and S14 *A* and *B*). The CHi-C analysis reveals an additional local interactive domain encompassing the marneral cluster that exhibits more pronounced contact intensities in roots where the cluster is transcribed, compared to leaves. This suggests a similar dual conformational switch between the actively transcribed and the silenced forms of the cluster as observed for the thalianol cluster (*SI Appendix*, Figs. S13 and S14 *C* and *D*). At the arabidiol/baruol cluster, we observed a single local interactive domain that precisely encompasses the cluster and shows increased contact frequency in leaves compared to roots and, therefore, negatively correlates with transcriptional activity (*SI Appendix*, Figs. S15 and S16).

Of note, the establishment of dynamic local 3D domains associated with clusters of coregulated genes may not be restricted to our target clusters. We observed strong interactions between a cluster of homologous ribulose biphosphate carboxylases genes encoding small subunits of the Rubisco enzyme and an adjacent region (48). These interactions are specific for roots where the cluster is silenced and are not detectable in leaves when the cluster is expressed (*SI Appendix*, Figs. S17 and S18). Future single-cell Hi-C analysis may further refine the correlation between chromosomal conformation of clusters and different expression states.

Cluster-Associated Silencing Domains Are Lost in the H3K27me3 Mutant. Our earlier work and studies in filamentous fungi have suggested that biosynthetic gene clusters are delineated by conserved chromatin modifications (20, 28, 49, 50). Among them is H3K27me3, a well-described histone mark primarily associated with gene silencing (51, 52). Peaks of H3K27me3 are detectable at all three metabolic gene clusters investigated here, and we have previously shown that cluster expression levels are elevated in *A. thaliana* mutant lines with reduced H3K27me3 levels (*SI Appendix*, Figs. S18–S20) (20, 52). Recent Hi-C analyses of *A. thaliana* chromosomes have described small interactive domains and chromatin loops that are enriched for H3K27me3 marks (38, 53). Therefore, we decided to reanalyze available Hi-C maps of H3K27me3 mutants and monitor the cluster-associated interactive domains (38).

For the thalianol, marneral, and arabidiol/baruol clusters, we found a significant reduction in the interaction strength of the associated interactive domains in H3K27me3-depleted chromatin suggesting a role for this chromatin mark in their formation (Fig. 4 *A–C* and *SI Appendix*, Fig. S22 *A–C*). In animals, a spatial domain that connects different gene clusters in 3D space has been shown to be constrained by H3K27me3 histone modifications (32, 34). Similarly, we detected high interaction counts between the thalianol and the marneral gene clusters, which are separated by 2.4 Mb on chromosome 5 (*SI Appendix*, Fig. S23*A*). The interaction intensity was strongly elevated in wild-type- versus the H3K27me3-depleted chromatin in reanalysis of existing Hi-C maps, thus, correlating with H3K27me3 levels (*SI Appendix*, Fig. S23*B*). In embryonic cells of mice, H3K27me3 labeled 3D domains are formed within central active nuclear regions in contrast to the peripheral localization of the H3K27me3 marked interactive domains identified here (34). In *Neurospora crassa*, loss of H3K27me3 marks in chromatin mutants leads to relocalization of subtelomeric regions toward the interior of the nucleus and upregulation of target genes (54).

Interestingly, at the cluster of homologous ribulose biphosphate carboxylases genes, we did not observe strong H3K27me3 markings in nonexpressing root tissues (*SI Appendix*, Fig. S24). In contrast, we identified significant cluster-associated H3K4me3 markings, well-described histone marks associated with gene activation, in the expressing leave tissues (*SI Appendix*, Fig. S24). We did not, however, detect strong enrichment of H3K4me3 markings at biosynthetic gene clusters (*SI Appendix*, Figs. S19–S21).

Metabolic Gene Clusters Reside in Local Interactive Domains in Diverse Plant Species. Next, we asked whether biosynthetic gene clusters are similarly located within distinct interactive domains in other plant species. To address this question, we analyzed available Hi-C maps of tomato, maize, and rice (41, 56). Each of these species contains, at least, one well-described metabolic gene cluster (*SI Appendix*, Fig. S25) (14). Chromosome topology in tomato, maize, and rice is characterized by a more pronounced structuring of

chromosomes into TADs compared to chromosome topology in *A. thaliana* (41, 56, 57). TADs are chromosomal regions with extensive internal chromatin contacts and limited interactions with adjacent regions (58, 59). As such, they resemble the cluster-associated interactive domains in *A. thaliana* described here. We, therefore, identified TADs in the respective genomes and analyzed their association with the maize DIMBOA and rice momilactone and phytocassane clusters as well as the tomatine biosynthetic gene cluster in tomato (Dataset S5) (17, 60–62). We observed that each cluster is positioned within a defined TAD that encompasses all individual cluster genes (Fig. 5 *A*–*D*). As seen for the biosynthetic gene clusters in *A. thaliana* and the ergot alkaloid

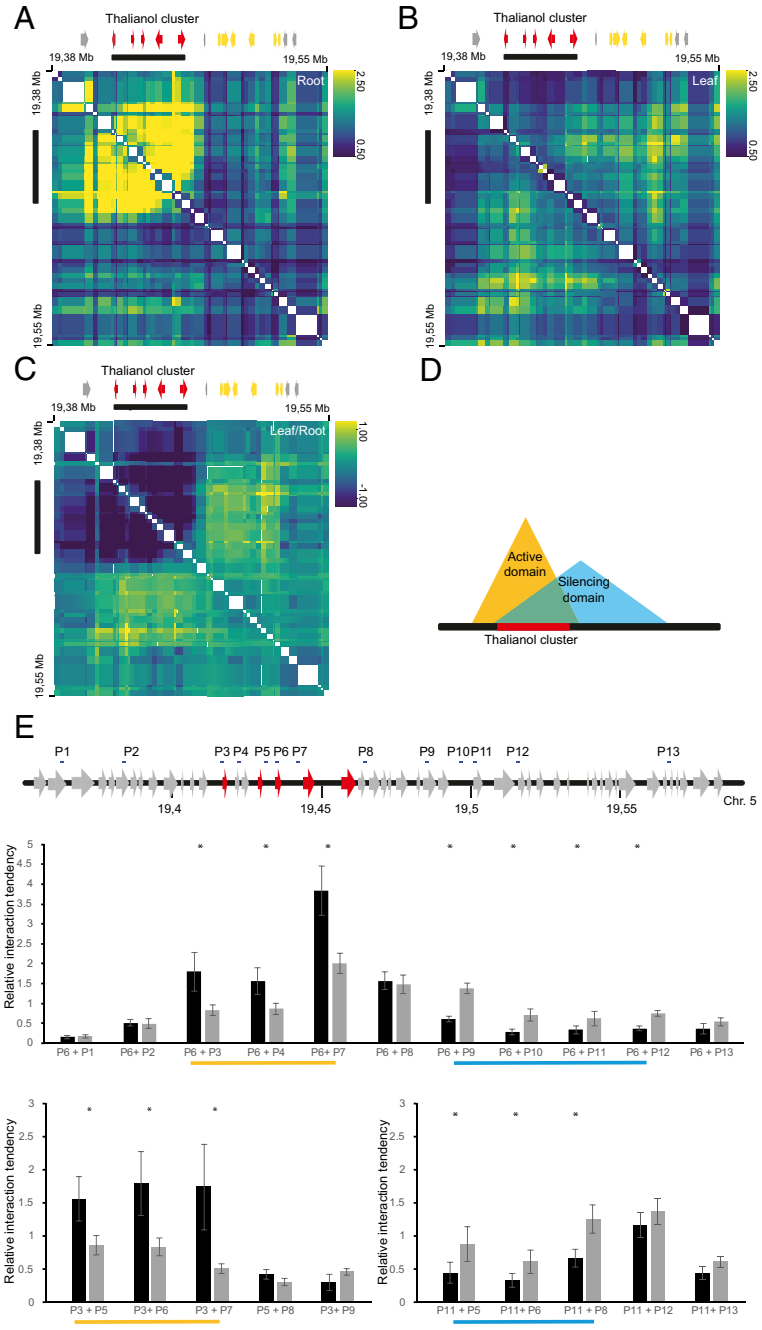

Fig. 3. Local interactive domains at the thalianol cluster. (*A* and *B*) The 2D CHi-C interaction map of the thalianol cluster region in roots (*A*) and leaves (*B*). Yellow to blue coloring indicates strong to weak interaction tendency. Genomic bin size: restriction fragment. (*C*) Differential 2D CHi-C interaction map of the thalianol cluster region in leaves versus roots. Coordinates are labeled from *Left* to *Right* and *Top* to *Bottom*. Cluster genes are shown in red, and additional genes with increased transcript levels in root versus leaves are shown in orange. Genes at the borders of the cluster-associated domain are in gray (from *Left* to *Right*—*At5g47910*, *At5g48050*, *At5g48150*, and *At5g48160*). Yellow to blue coloring indicates strong to weak interaction tendency. Genomic bin size: restriction fragment. (*D*) Model of interactive domains at the thalianol cluster. A strong interactive domain is formed at the thalianol cluster during active transcription, and a weaker but larger domain is formed during transcriptional repression. (*E*) The 3C analysis of the extended thalianol cluster region. *Top*, the chromosomal area around the thalianol cluster. Cluster genes are shown in red, and noncluster genes are shown in gray. The histograms below display the 3C interaction profile of different sites in the extended cluster region. Interaction tendencies in roots and leaves are shown in black and gray bars, respectively. Significantly increased interactions in roots are underlined in yellow, and significantly increased interactions in leaves are underlined in blue. The asterisks indicate a significant difference between interaction tendencies (Student *t* test, $P < 0.01$).

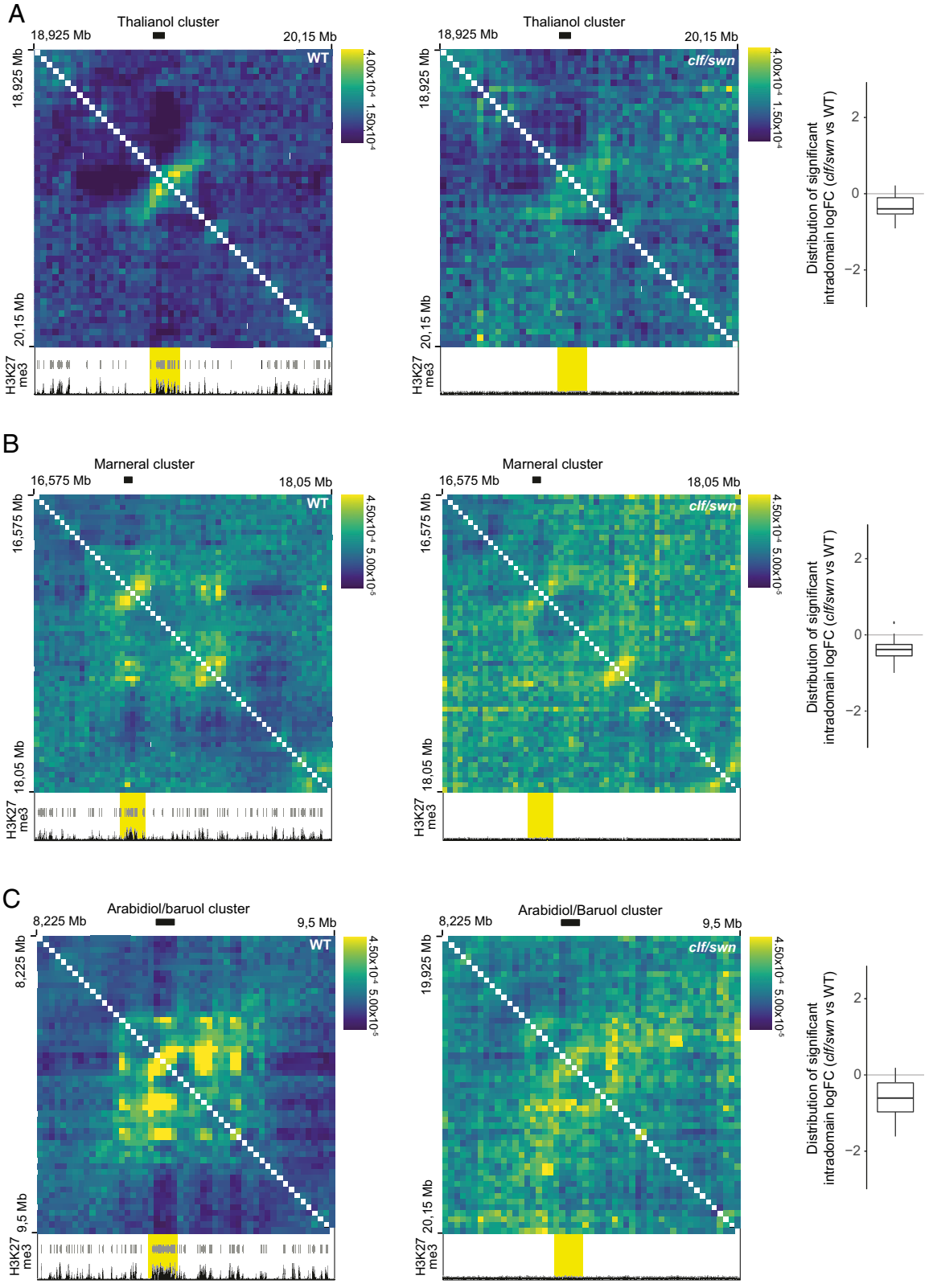

Fig. 4. Loss of cluster-associated interactive domains in a H3K27me3 mutant. (*A*–*C*) 2D Hi-C interaction maps of the 3D conformation surrounding the thalianol cluster (*A*), marneral (*B*), and arabidiol/baruol (*C*) clusters in seedlings. *Left*, wild type; *Right clf/swn* mutant. Chromosomes are labeled from *Left* to *Right* and *Top* to *Bottom*. Yellow to blue coloring indicates strong to weak interaction tendency. Genomic bin size: 25 kb. The boxplots on the right show logarithmic fold change values for the wild-type versus the *clf/swn* mutant intradomain interaction counts. The tracks at the *Bottom* of each panel show significant peaks and enrichment tracks of H3K27me3 markings in wild-type (*Left*) and *clf/swn* double mutants extracted from Shu et al. (55). In yellow, area of cluster-associated interactive domain.

EAS cluster of *E. festucae*, cluster-associated TADs include additional genes outside the respective clusters (Fig. 5 *A–D*) (35). We did not observe obvious long-range contacts with other regions of the genome for the investigated clusters. TADs in maize, rice, and tomato are suggested to be separated by different expression and epigenetic states but are not associated with coexpression of the genes located within the same domain (41). All four clusters investigated here show strong coexpression pattern, and for clusters in rice and maize, we have previously reported an enrichment for H3K27me3 chromatin modifications (20, 63). Analysis of gene expression datasets associated with the investigated Hi-C maps show high expression levels for the tomatine cluster, medium expression levels for the DIMBOA cluster, and very low expression levels for both rice clusters (*SI Appendix*, Fig. S26 *A–D*). This suggests that formation of cluster-associated domains is not restricted to specific expression states. Future comparative studies analyzing chromosome architecture under conditions with variable cluster expression states in maize, rice, and tomato should shed light on the structural flexibility of the cluster-encompassing TADs in these other species.

Conclusions

To summarize, we report that metabolic gene clusters reside in defined local interactive domains in plant genomes. In *A. thaliana*,

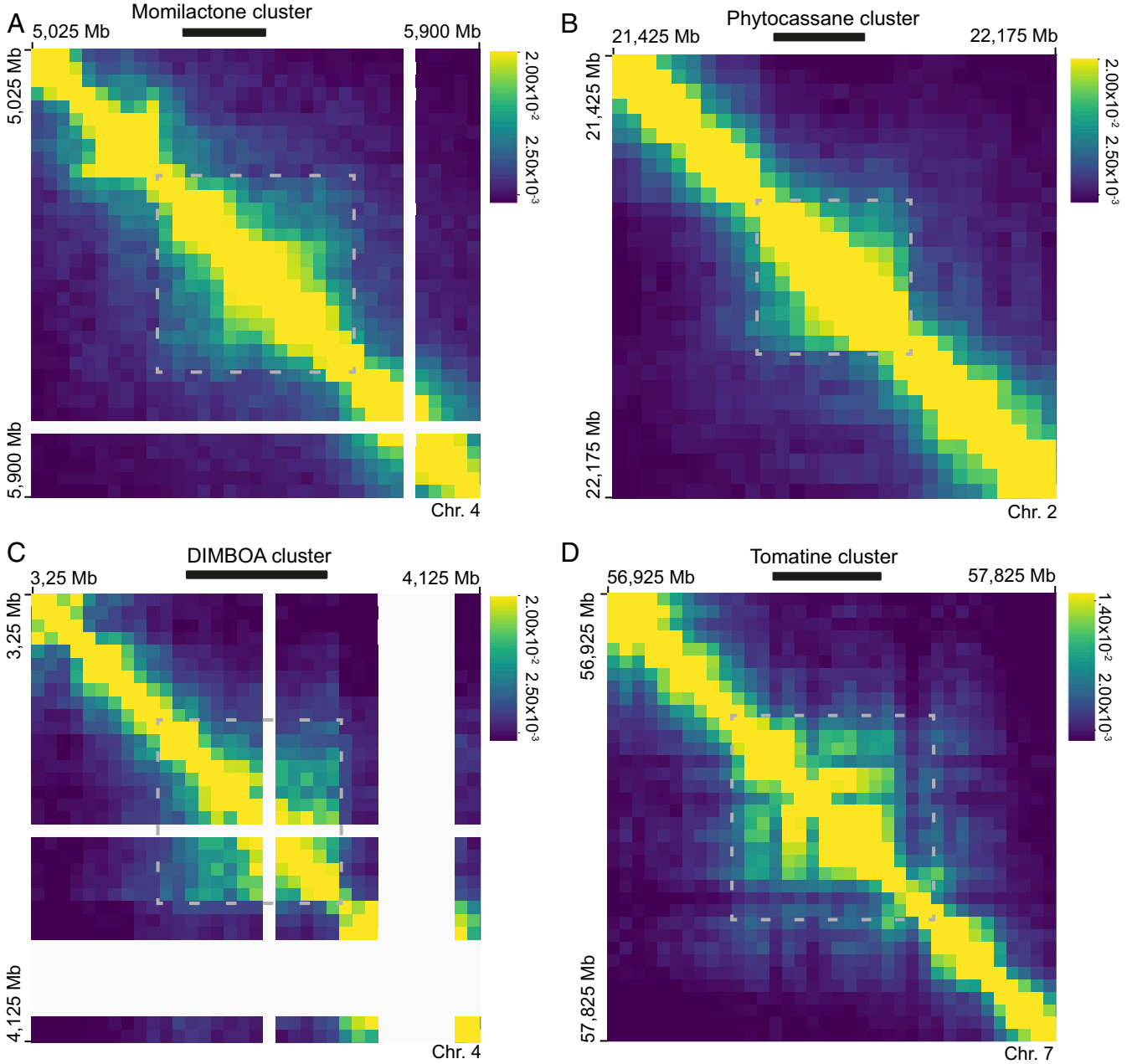

Fig. 5. Metabolic gene clusters from other plant species are located in single TADs. (*A–D*) The 2D Hi-C interaction maps of the 3D conformations of the momilactone (*A*) and phytocassane (*B*) cluster-associated genomic regions in rice; the DIMBOA cluster-associated genomic region in maize (*C*); and the tomatine cluster-associated genomic region in tomato (*D*). Chromosomes are labeled from *Left* to *Right* and *Top* to *Bottom*. The biosynthetic gene clusters are indicated with a black line, and associated TADs are shown in the gray boxes. Yellow to blue coloring indicates strong to weak interaction tendency. Genomic bin size: 25 kb.

the structure of these domains is flexible and changes its configuration between organs that express or do not express the cluster in question.

Local interactive domains surrounding clusters of genes may insulate these clusters from their chromosomal neighborhood. This may prevent the spreading of repressive and active chromatin environments at gene clusters into nearby chromosome areas (64, 65). Cluster-associated interactive domains may further constitute local microenvironments that support tight coregulation of gene expression. Local domains with increased internal contacts have been observed for neighboring functionally unrelated genes with shared transcriptional and epigenetic states in different eukaryotic species (66, 67). The formation of these domains may indicate a general principle in eukaryotic genome organization and may provide a structural platform for evolution of functionally related and coordinately regulated gene clusters.

Structural flexibility of 3D domains at biosynthetic gene clusters is accompanied by repositioning of clusters inside the nuclear space. Peripheral localization of clusters is observed in nonexpressing organs and interior localization in expressing organs (*SI Appendix*, Fig. S27). As such, these cluster regions are among the most dynamic regions in the *A. thaliana* genome. The relocation of clusters toward heterochromatic areas of the nucleus may be important in the efficient silencing of these clusters. It has been shown that misexpression of thalianol cluster genes leads to severe developmental defects in the plant, and so, tight coordinate regulation of the thalianol cluster genes is likely to be critical for survival (24). The colocalization of pathway genes in operon-like gene clusters may directly support the formation of single 3D domains. Such conformation may, in turn, facilitate coordinate engagement of these genes in nuclear repositioning and transcriptional coregulation as compared to a scenario in which genes are dispersed in different chromosomal locations and embedded in separate 3D domains.

Furthermore, we show that loss of the histone mark H3K27me3 is associated with alterations in the interaction intensity of the cluster-associated chromosome architecture, supporting the important role for this chromatin modification in shaping chromosome structure in *A. thaliana* (38, 68).

Our results reveal the complex chromosomal architecture surrounding metabolic gene clusters and shed light on the potential mechanisms that constrain coexpression within clusters of eukaryotic genes. We show that clustering of genes on the linear eukaryotic chromosome is accompanied by compartmentalisation of the 3D chromosome. Furthermore, we provide evidence that the spatial organization of plant chromosomes is plastic. These advances will provide the basis for future studies to better understand the role of chromosome organization in defining gene cluster structure, expression, and evolution.

Materials and Methods

A. thaliana plants used in this study were of the Col-0 wild type. For all experiments, *A. thaliana* seeds were surface sterilized and grown vertically on Petri dishes containing Murashige and Skoog plant salt medium supplemented with 0.5% phytagel and 0.75% sucrose (69). Plants were grown at 22 °C with a 16-h light/8-h dark photoperiod for 7 d. Triplicate 3C, duplicate Hi-C and CHi-C, and triplicate RNA-seq experiments were performed as described in *SI Appendix, SI Materials and Methods*. DNA FISH experiments were essentially performed as described before and are outlined in *SI Appendix, SI Materials and Methods* (70, 71). All Hi-C, CHi-C, and RNA-seq data were deposited in the National Center for Biotechnology Information (NCBI) Gene Expression Omnibus (GEO) database (accession no. PRJNA576277) (72). *A. thaliana* wild-type and H3K27me3 mutant Hi-C datasets (accession no. SRP043612), ChIPseq datasets (GSE108960), as well as tomato, rice, and maize Hi-C datasets (PRJNA486213 and PRJNA391551) were previously reported (38, 41, 55, 56, 73–76).

ACKNOWLEDGMENTS. We thank Stefan Schoenfelder and Simon Andrews for helpful discussions in setting up CHi-C protocols and defining capture probes, Silin Zhong for directing us to the Hi-C-associated RNAseq datasets, and Kasia Oktaba for a discussion on Hi-C experimental details. This work was supported by the Royal Society funded University Research Fellowship UF160138 (H.-W.N.), Newton Advanced Fellowship NAF\R1\180303 (S.L.F.-V., L.H.), a Marie Curie Actions EMBO Long-Term Fellowship (H.-W.N.), the joint Engineering and Physical Sciences Research Council/BBSRC-funded OpenPlant Synthetic Biology Research Centre Grant BB/L014130/1 (H.-W.N., A.O.), CONACYT Masters Scholarships (A.R.-C., J.E.S.-F.), CONACYT Research Fellowship 2015 (S.L.F.-V.), the UK Biotechnological and Biological Sciences Research Council Institute Strategic Programme Grants "Molecules from Nature" BBS/E/J/000PR9790 (A.O.) and BB/J004480/1 (P.F., S.W.W.), the John Innes Foundation (A.O.), and the University of Bath (H.-W.N., L.H.).

1. E. P. Rocha, The organization of the bacterial genome. *Annu. Rev. Genet.* **42**, 211–233 (2008).
2. C. Pál, L. D. Hurst, Evidence for co-evolution of gene order and recombination rate. *Nat. Genet.* **33**, 392–395 (2003).
3. L. D. Hurst, C. Pál, M. J. Lercher, The evolutionary dynamics of eukaryotic gene order. *Nat. Rev. Genet.* **5**, 299–310 (2004).
4. P. Michalak, Coexpression, coregulation, and cofunctionality of neighboring genes in eukaryotic genomes. *Genomics* **91**, 243–248 (2008).
5. A. E. Osbourn, B. Field, Operons. *Cell. Mol. Life Sci.* **66**, 3755–3775 (2009).
6. A. Rokas, J. H. Wisecaver, A. L. Lind, The birth, evolution and death of metabolic gene clusters in fungi. *Nat. Rev. Microbiol.* **16**, 731–744 (2018).
7. H. W. Nützmann, C. Scazzocchio, A. Osbourn, Metabolic gene clusters in eukaryotes. *Annu. Rev. Genet.* **52**, 159–183 (2018).
8. J. C. Slot, Fungal gene cluster diversity and evolution. *Adv. Genet.* **100**, 141–178 (2017).
9. N. H. Giles *et al*., Gene organization and regulation in the qa (quinic acid) gene cluster of *Neurospora crassa*. *Microbiol. Rev.* **49**, 338–358 (1985).
10. S. Wong, K. H. Wolfe, Birth of a metabolic gene cluster in yeast by adaptive gene relocation. *Nat. Genet.* **37**, 777–782 (2005).
11. J. D. Walton, Horizontal gene transfer and the evolution of secondary metabolite gene clusters in fungi: An hypothesis. *Fungal Genet. Biol.* **30**, 167–171 (2000).
12. D. Sproul, N. Gilbert, W. A. Bickmore, The role of chromatin structure in regulating the expression of clustered genes. *Nat. Rev. Genet.* **6**, 775–781 (2005).
13. A. C. C. Huang *et al*., A specialized metabolic network selectively modulates *Arabidopsis* root microbiota. *Science* **364**, eaau6389 (2019).
14. H. W. Nützmann, A. Huang, A. Osbourn, Plant metabolic clusters - from genetics to genomics. *New Phytol.* **211**, 771–789 (2016).
15. L. M. Schneider *et al*., The Cer-cqu gene cluster determines three key players in a β-diketone synthase polyketide pathway synthesizing aliphatics in epicuticular waxes. *J. Exp. Bot.* **68**, 5009 (2017).
16. T. Winzer *et al*., A *Papaver somniferum* 10-gene cluster for synthesis of the anticancer alkaloid noscapine. *Science* **336**, 1704–1708 (2012).
17. M. Itkin *et al*., Biosynthesis of antinutritional alkaloids in solanaceous crops is mediated by clustered genes. *Science* **341**, 175–179 (2013).
18. S. A. Kautsar, H. G. Suarez Duran, K. Blin, A. Osbourn, M. H. Medema, plantiSMASH: automated identification, annotation and expression analysis of plant biosynthetic gene clusters. *Nucleic Acids Res.* **45**, W55–W63 (2017).
19. Y. Shang *et al*., Plant science. Biosynthesis, regulation, and domestication of bitterness in cucumber. *Science* **346**, 1084–1088 (2014).
20. N. Yu *et al*., Delineation of metabolic gene clusters in plant genomes by chromatin signatures. *Nucleic Acids Res.* **44**, 2255–2265 (2016).
21. A. M. Boutanaev *et al*., Investigation of terpene diversification across multiple sequenced plant genomes. *Proc. Natl. Acad. Sci. U.S.A.* **112**, E81–E88 (2015).
22. A. M. Boutanaev, A. E. Osbourn, Multigenome analysis implicates miniature inverted-repeat transposable elements (MITEs) in metabolic diversification in eudicots. *Proc. Natl. Acad. Sci. U.S.A.* **115**, E6650–E6658 (2018).
23. B. Field *et al*., Formation of plant metabolic gene clusters within dynamic chromosomal regions. *Proc. Natl. Acad. Sci. U.S.A.* **108**, 16116–16121 (2011).
24. B. Field, A. E. Osbourn, Metabolic diversification—independent assembly of operon-like gene clusters in different plants. *Science* **320**, 543–547 (2008).
25. E. Wegel, R. Koumproglou, P. Shaw, A. Osbourn, Cell type-specific chromatin decondensation of a metabolic gene cluster in oats. *Plant Cell* **21**, 3926–3936 (2009).
26. D. P. Barlow, Genomic imprinting: A mammalian epigenetic discovery model. *Annu. Rev. Genet.* **45**, 379–403 (2011).
27. P. D. Cárdenas *et al*., GAME9 regulates the biosynthesis of steroidal alkaloids and upstream isoprenoids in the plant mevalonate pathway. *Nat. Commun.* **7**, 10654 (2016).
28. A. Gacek, J. Strauss, The chromatin code of fungal secondary metabolite gene clusters. *Appl. Microbiol. Biotechnol.* **95**, 1389–1404 (2012).
29. D. Lohr, P. Venkov, J. Zlatanova, Transcriptional regulation in the yeast *GAL* gene family: A complex genetic network. *FASEB J.* **9**, 777–787 (1995).
30. H. W. Nützmann, A. Osbourn, Gene clustering in plant specialized metabolism. *Curr. Opin. Biotechnol.* **26**, 91–99 (2014).
31. T. Ragoczy, M. A. Bender, A. Telling, R. Byron, M. Groudine, The locus control region is required for association of the murine β-globin locus with engaged transcription factories during erythroid maturation. *Genes Dev.* **20**, 1447–1457 (2006).

32. S. Schoenfelder *et al*., Polycomb repressive complex PRC1 spatially constrains the mouse embryonic stem cell genome. *Nat. Genet.* **47**, 1179–1186 (2015).
33. M. Simonis *et al*., Nuclear organization of active and inactive chromatin domains uncovered by chromosome conformation capture-on-chip (4C). *Nat. Genet.* **38**, 1348–1354 (2006).
34. M. Vieux-Rochas, P. J. Fabre, M. Leleu, D. Duboule, D. Noordermeer, Clustering of mammalian Hox genes with other H3K27me3 targets within an active nuclear domain. *Proc. Natl. Acad. Sci. U.S.A.* **112**, 4672–4677 (2015).
35. D. J. Winter *et al*., Repeat elements organise 3D genome structure and mediate transcription in the filamentous fungus *Epichloë festucae*. *PLoS Genet.* **14**, e1007467 (2018).
36. R. Sohrabi *et al*., In planta variation of volatile biosynthesis: An alternative biosynthetic route to the formation of the pathogen-induced volatile homoterpene DMNT via triterpene degradation in *Arabidopsis* roots. *Plant Cell* **27**, 874–890 (2015).
37. H. W. Nützmann, A. Osbourn, Regulation of metabolic gene clusters in *Arabidopsis thaliana*. *New Phytol.* **205**, 503–510 (2015).
38. S. Feng *et al*., Genome-wide Hi-C analyses in wild-type and mutants reveal high-resolution chromatin interactions in *Arabidopsis*. *Mol. Cell* **55**, 694–707 (2014).
39. S. Grob, M. W. Schmid, U. Grossniklaus, Hi-C analysis in *Arabidopsis* identifies the KNOT, a structure with similarities to the flamenco locus of *Drosophila*. *Mol. Cell* **55**, 678–693 (2014).
40. C. Wang *et al*., Genome-wide analysis of local chromatin packing in *Arabidopsis thaliana*. *Genome Res.* **25**, 246–256 (2015).
41. P. Dong *et al*., Tissue-specific Hi-C analyses of rice, foxtail millet and maize suggest non-canonical function of plant chromatin domains. *J. Integr. Plant Biol.* **62**, 201–217 (2020).
42. S. Del Prete, J. Arpón, K. Sakai, P. Andrey, V. Gaudin, Nuclear architecture and chromatin dynamics in interphase nuclei of *Arabidopsis thaliana*. *Cytogenet. Genome Res.* **143**, 28–50 (2014).
43. P. Fransz, J. H. De Jong, M. Lysak, M. R. Castiglione, I. Schubert, Interphase chromosomes in *Arabidopsis* are organized as well defined chromocenters from which euchromatin loops emanate. *Proc. Natl. Acad. Sci. U.S.A.* **99**, 14584–14589 (2002).
44. L. Simon, M. Voisin, C. Tatout, A. V. Probst, Structure and function of centromeric and pericentromeric heterochromatin in *Arabidopsis thaliana*. *Front Plant Sci* **6**, 1049 (2015).
45. P. Martin *et al*., Capture Hi-C reveals novel candidate genes and complex long-range interactions with related autoimmune risk loci. *Nat. Commun.* **6**, 10069 (2015).
46. B. Mifsud *et al*., Mapping long-range promoter contacts in human cells with high-resolution capture Hi-C. *Nat. Genet.* **47**, 598–606 (2015).
47. G. Andrey *et al*., A switch between topological domains underlies HoxD genes collinearity in mouse limbs. *Science* **340**, 1234167 (2013).
48. M. Izumi, H. Tsunoda, Y. Suzuki, A. Makino, H. Ishida, RBCS1A and RBCS3B, two major members within the *Arabidopsis* RBCS multigene family, function to yield sufficient Rubisco content for leaf photosynthetic capacity. *J. Exp. Bot.* **63**, 2159–2170 (2012).
49. L. R. Connolly, K. M. Smith, M. Freitag, The *Fusarium graminearum* histone H3 K27 methyltransferase KMT6 regulates development and expression of secondary metabolite gene clusters. *PLoS Genet.* **9**, e1003916 (2013).
50. L. Studt *et al*., Knock-down of the methyltransferase Kmt6 relieves H3K27me3 and results in induction of cryptic and otherwise silent secondary metabolite gene clusters in *Fusarium fujikuroi*. *Environ. Microbiol.* **18**, 4037–4054 (2016).
51. S. Feng, S. E. Jacobsen, Epigenetic modifications in plants: An evolutionary perspective. *Curr. Opin. Plant Biol.* **14**, 179–186 (2011).
52. X. Zhang *et al*., Whole-genome analysis of histone H3 lysine 27 trimethylation in *Arabidopsis*. *PLoS Biol.* **5**, e129 (2007).
53. C. Liu *et al*., Genome-wide analysis of chromatin packing in *Arabidopsis thaliana* at single-gene resolution. *Genome Res.* **26**, 1057–1068 (2016).
54. A. D. Klocko *et al*., Normal chromosome conformation depends on subtelomeric facultative heterochromatin in *Neurospora crassa*. *Proc. Natl. Acad. Sci. U.S.A.* **113**, 15048–15053 (2016).
55. J. Shu *et al*., Genome-wide occupancy of histone H3K27 methyltransferases CURLY LEAF and SWINGER in *Arabidopsis* seedlings *Plant Direct*, 10.1002/pld3.100 (2019).
56. P. Dong *et al*., 3D chromatin architecture of large plant genomes determined by local A/B compartments. *Mol. Plant* **10**, 1497–1509 (2017).
57. C. Liu, Y. J. Cheng, J. W. Wang, D. Weigel, Prominent topologically associated domains differentiate global chromatin packing in rice from *Arabidopsis*. *Nat. Plants* **3**, 742–748 (2017).
58. J. R. Dixon *et al*., Topological domains in mammalian genomes identified by analysis of chromatin interactions. *Nature* **485**, 376–380 (2012).
59. T. Sexton *et al*., Three-dimensional folding and functional organization principles of the Drosophila genome. *Cell* **148**, 458–472 (2012).
60. M. Frey *et al*., Analysis of a chemical plant defense mechanism in grasses. *Science* **277**, 696–699 (1997).
61. S. Swaminathan, D. Morrone, Q. Wang, D. B. Fulton, R. J. Peters, CYP76M7 is an *ent*-cassadiene C11α-hydroxylase defining a second multifunctional diterpenoid biosynthetic gene cluster in rice. *Plant Cell* **21**, 3315–3325 (2009).
62. P. R. Wilderman, M. Xu, Y. Jin, R. M. Coates, R. J. Peters, 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 (2004).
63. J. H. Wisecaver *et al*., A global coexpression network approach for connecting genes to specialized metabolic pathways in plants. *Plant Cell* **29**, 944–959 (2017).
64. R. D. Acemel, I. Maeso, J. L. Gómez-Skarmeta, Topologically associated domains: A successful scaffold for the evolution of gene regulation in animals. *Wiley Interdiscip. Rev. Dev. Biol.* **6**, e265 (2017).
65. J. R. Dixon, D. U. Gorkin, B. Ren, Chromatin domains: The unit of chromosome organization. *Mol. Cell* **62**, 668–680 (2016).
66. S. S. P. Rao *et al*., Cohesin loss eliminates all loop domains. *Cell* **171**, 305–320.e24 (2017).
67. M. J. Rowley *et al*., Evolutionarily conserved principles predict 3D chromatin organization. *Mol. Cell* **67**, 837–852.e7 (2017).
68. W. Zhu *et al*., Altered chromatin compaction and histone methylation drive non-additive gene expression in an interspecific *Arabidopsis* hybrid. *Genome Biol.* **18**, 157 (2017).
69. T. Murashige, F. Skoog, A revised medium for rapid growth and bio assays with tobacco tissue cultures. *Physiol. Plant.* **15**, 473–497 (1962).
70. A. C. Martín, P. Shaw, D. Phillips, S. Reader, G. Moore, Licensing MLH1 sites for crossover during meiosis. *Nat. Commun.* **5**, 4580 (2014).
71. A. Pendle, P. Shaw, Immunolabeling and in situ labeling of isolated plant interphase nuclei. *Methods Mol. Biol.* **1429**, 65–76 (2016).
72. D. Doerr, H. W. Nützmann, Chromosome structure at metabolic gene clusters BioProject. https://www.ncbi.nlm.nih.gov/bioproject/PRJNA576277. Deposited 7 October 2019.
73. UCLA-NL, Arabidopsis thaliana clf-28 swn-7 for Hi-C (HindIII). Sequence Read Archive. https://www.ncbi.nlm.nih.gov/sra/SRP043612. Accessed 11 March 2019.
74. J. Shu, Y. Cui, Genome-wide occupancy of histone H3K27 methyltransferases CURLY LEAF and SWINGER in Arabidopsis seedlings. Gene Expression Omnibus. https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE108960. Accessed 28 February 2020.
75. The Chinese University of Hong Kong, C3C4 ENCODE project BioProject. https://www.ncbi.nlm.nih.gov/bioproject/PRJNA486213. Accessed 20 May 2019.
76. The Chinese University of Hong Kong, Diverse 3D chromatin architecture of medium and large plant genomes. BioProject. https://www.ncbi.nlm.nih.gov/bioproject/PRJNA391551. Accessed 14 February 2020.