

Order in the playground

Formation of plant gene clusters in dynamic chromosomal regions

Ben Field^{1,*} and Anne Osbourn²

¹Laboratoire de Génétique et de Biophysique des Plantes (LGBP); CNRS-CEA-Université de la Méditerranée; Marseille, France; ²Department of Metabolic Biology; John Innes Centre; Norwich Research Park; Norwich UK

In eukaryotes the arrangement of genes along the chromosome is not as random as it at first appeared, and distinctive clusters of functionally related but non-homologous genes can be found in the genomes of certain animals and fungi. These include the major histocompatibility complex in mammals and gene clusters for nutrient use and secondary metabolite production in fungi. A growing number of functional gene clusters for different types of secondary metabolite are now being discovered in plant genomes. However, the molecular mechanisms and evolutionary pressures behind their formation are poorly understood. Here we discuss the implications of our recent investigation into the origin of two functional gene clusters in the model plant *Arabidopsis thaliana*.

We recently reported the discovery of the marneral cluster, a functional gene cluster for the biosynthesis of previously unknown triterpene derivatives, in the model plant *Arabidopsis thaliana*.¹ Our study is one of the latest of a growing number of reports on the identification of functional gene clusters in plants. In eukaryotes functionally related genes are usually scattered across the genome. However, clusters of functionally related but non-homologous genes have been identified in the genomes of animals and certain fungi.^{2,3} These include the MHC in mammals,⁴ gene clusters for nutrient use in yeast,⁵⁻⁷ and a multitude of clusters for diverse secondary metabolic pathways in filamentous fungi.^{8,9} Although these

clusters have operon-like features (physical clustering and co-regulation) they are clearly distinct from operons because the genes within each cluster are independently transcribed.³

In plants, genes for secondary metabolic pathways are not generally considered to be clustered. For example, genes for the well studied phenylpropanoid and glucosinolate pathways are not clustered in the model plant *Arabidopsis thaliana*. However, as more plant genomes are sequenced and analyzed it is becoming clear that clustering of genes for secondary metabolism may be more common than previously realized. Ten years ago a gene cluster for the synthesis of cyclic hydroxamic acids associated with plant defense was discovered in maize (*Zea mays*).¹⁰ Since then gene clusters for defense-related secondary metabolic pathways have been discovered in a variety of other species, including oat (*Avena strigosa*),¹¹⁻¹⁴ rice (*Oryza sativa*),¹⁵⁻¹⁸ sorghum (*Sorghum bicolor*),¹⁹ cassava (*Manihot esculenta*),¹⁹ *Lotus japonicas*¹⁹ and *Arabidopsis thaliana*.^{1,3} The discovery of these functional gene clusters in plants, with other as yet unpublished examples in the pipeline, has allowed us to start looking for the rules that may underlie gene cluster formation and genome dynamics in plants.

Two functional gene clusters have been identified in the genome of *Arabidopsis thaliana* so far- first the thalianol cluster, and more recently the marneral cluster (Fig. 1).^{1,3} Each cluster consists of a set of co-expressed genes that encode enzymatic pathways for the synthesis and elaboration of either thalianol or marnerol, specialized

Keywords: secondary metabolism, triterpene, *Arabidopsis*, gene cluster, genome evolution

Submitted: 11/11/11

Accepted: 01/13/12

<http://dx.doi.org/10.4161/mge.19348>

*Correspondence to: Ben Field;
Email: ben.field@univmed.fr

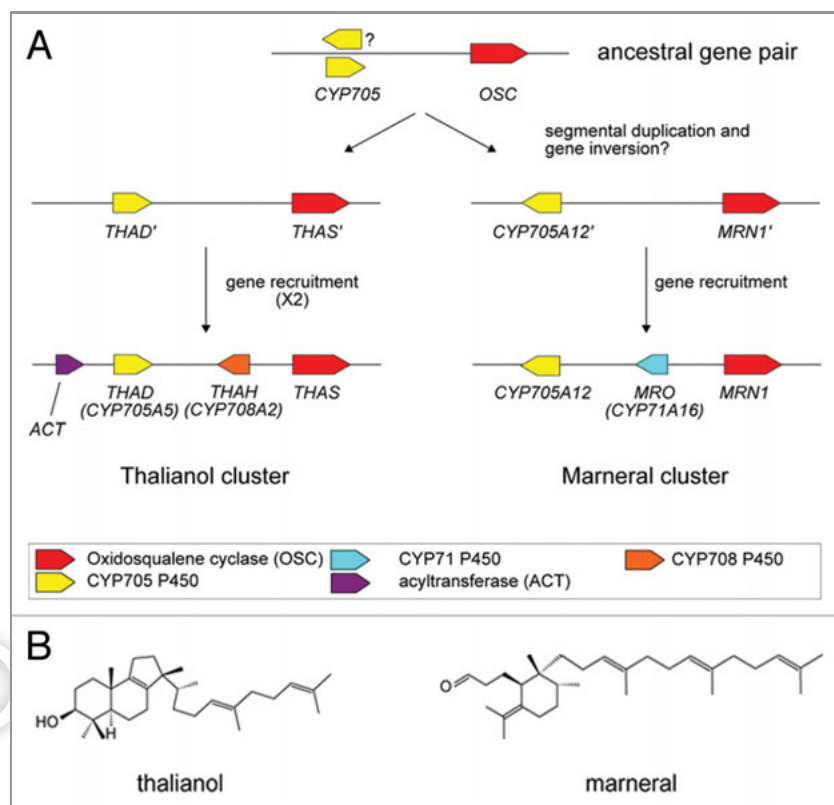


Figure 1. The marneral and thalianol gene clusters in *A. thaliana*. (A) The present day thalianol and marneral clusters (bottom) and a proposed scheme for their formation based on the assumption that the two clusters were founded by duplication of an ancestral *OSC/CYP705* gene pair. (B) Structures of thalianol, synthesized by THAS in the first committed step of the thalianol pathway; and marneral, synthesized by MRN1 in the first committed step of the marneral pathway. Figure adapted from reference 1.

tricyclic triterpenes that so far have only been identified in *A. thaliana* (2,19). The functions of these elaborated triterpenes are not yet known, although triterpene derivatives are often involved in defense. A possible role for thalianol or thalianol derivatives in abiotic stress responses has also recently emerged from a study where a calcium-dependent lipid-binding protein was shown to control the expression of *THAS* in the thalianol cluster.²⁰ The discovery of two functional gene clusters within the sequenced *A. thaliana* genome and the availability of genome sequences for other related species such as *Arabidopsis lyrata*, *Brassica rapa*, *Thellungiella parvula* and *Thellungiella halophila* has provided the opportunity to use comparative genomics approaches to understand cluster formation. A discussion of our endeavors in this area, with additional insights from new published findings, will be the focus of this article.

The Timing of Gene Cluster Assembly

Analysis of the *Arabidopsis lyrata* genome (a close relative of *A. thaliana*) revealed that the thalianol cluster is present, and that the marneral cluster is absent.¹ Two observations suggest that the absence of the marneral cluster is due to a deletion event. First, we found that two of the marneral cluster gene products (MRN1 and CYP705A12) have more basal phylogenetic origins than their thalianol cluster paralog (THAS and THAD). Second, the thalianol cluster genes from *A. thaliana* each have a direct ortholog in *A. lyrata*, while the marneral cluster genes do not. This evidence combined with phylogenetic evidence for the individual genes within each cluster allowed us to propose that gene cluster assembly occurred before the divergence of *A. lyrata* and *A. thaliana*, or about 13 million years ago based on a

dated phylogeny.²¹ Our phylogenetic evidence also showed that the marneral and thalianol cluster genes are restricted to the Brassicaceae lineage, putting an upper limit on gene cluster age. More recent evidence now suggests that the thalianol and marneral clusters are likely to be restricted to a single clade within the Brassicaceae: we cannot detect intact clusters in the newly sequenced genomes of *Brassica rapa*, *Thellungiella parvula* or *Thellungiella halophila*.²²⁻²⁴ Interestingly there are also no direct orthologs to the majority of the cluster genes. These findings suggest that the clusters and the majority of associated genes were not yet present in the common ancestor of *B. rapa*/*Thellungiella* and *A. thaliana*, which is thought to have existed about 43 million years ago²¹ (Fig. 2). An alternative, although less-parsimonious, explanation is that the gene clusters have been lost from both *B. rapa* and *Thellungiella*. This explanation would place cluster formation considerably earlier at between 43 and 64 million years ago, or around the same time as the α -whole genome duplication event. Whole genome sequencing of more Brassicaceae species, such as *Capsella bursa* which is currently underway, will increase our resolution of the timing of cluster formation.

Location Matters

The thalianol and marneral clusters are not simply the result of the whole-scale duplication of an ancestral cluster. Instead, the two clusters were most probably founded either independently, or else by the duplication of an ancestral gene pair followed by independent rearrangements and the recruitment of different genes (Fig. 1). Crucially, these hypotheses suggest that the marneral and thalianol cluster regions could share special features that are required for cluster formation.

Each cluster is located in an island between segments of the genome that were duplicated in the most recent whole genome duplication event, which occurred ~47–69 mya (the α -event). This location is rather unusual, because only 11% of annotated genes exist outside of α -duplicated segments. The discovery of both gene clusters in non- α regions implies that such locations may be

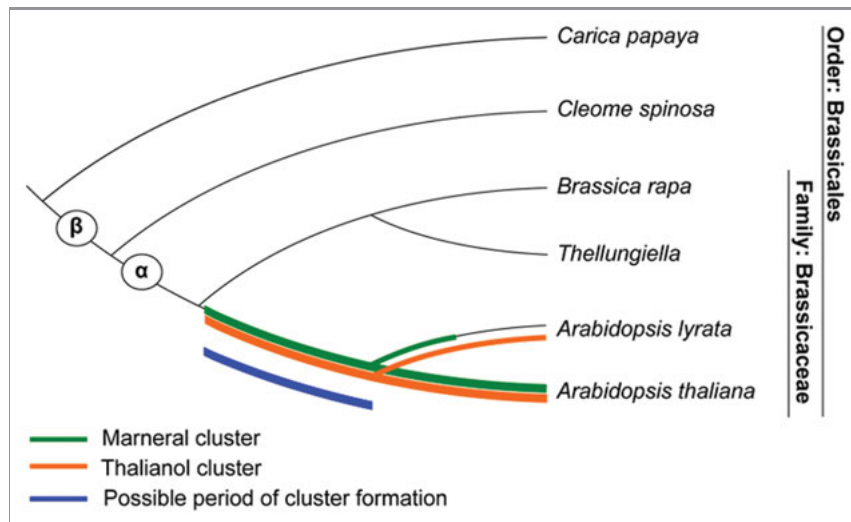


Figure 2. Timing of cluster assembly. The evolutionary tree highlights the period over which the thalianol (orange line) and marneral (green line) clusters may have been present in the Brassicaceae. The blue line indicates the period over which the thalianol and marneral clusters probably formed. These timings are predicted by the most parsimonious explanation of gene cluster assembly (see text). The α and β whole genome duplication events are indicated as circles. Figure adapted from reference 1.

important for gene cluster assembly. Non- α regions are known to be generally poor in genes and to contain an excess of TEs and pseudogenes.^{25,26} We found that TE density was indeed higher for the cluster regions than in flanking chromosomal regions. Cluster TE density was also higher than in other comparable regions between α -duplicated chromosome segments. It is tempting to suggest that these TEs directly contributed to the assembly of gene clusters: TEs promote ectopic recombination, and certain classes such as helitrons, which are present in both clusters, can transduplicate genes.^{27,28} However, due to their relatively short half lives (0.6–2 million years in *Arabidopsis*²⁹) and an almost complete absence of TE conservation between *A. thaliana* and *A. lyrata* it is very unlikely that present day TEs directly contributed to gene cluster assembly. At the same time, the short lifespan of TEs means that they serve as excellent markers of recent ectopic recombination activity. The TE enrichment at the marneral and thalianol clusters therefore suggests that these regions are particularly dynamic because they accept and retain segmental duplications (SDs) more frequently than neighboring regions. SD acceptor sites have been identified in other eukaryotic genomes and are, for

example, present in the human genome where they are thought to play a key role in the generation of new gene functions.³⁰ Interestingly, human SD acceptor sites are

suggested to have arisen following a burst in retrotransposition activity (the *Alu* burst) 35 million years ago, implicating TEs in the establishment of SD acceptor regions.³⁰

Is enhanced recombinational activity an inherent feature of the cluster regions? The evidence appears to indicate that it is. The thalianol cluster is TE rich and has a completely different TE composition in both *A. thaliana* and in *A. lyrata*. This suggests that the thalianol region was still a recombinational hotspot in the *lyrata*/*thaliana* ancestor, and that recombinational activity is independent of the TE composition of these regions. The thalianol and marneral clusters also show evidence of large-scale recombination events. In *A. lyrata* the marneral cluster appears to have been entirely deleted, and the thalianol cluster has experienced the insertion of a large stretch of TE rich DNA.³¹ A synteny map between *A. thaliana* and *A. lyrata* shows that the two cluster regions are located either side of the centromere of *A. lyrata* chromosome 8 due to chromosomal rearrangements (Fig. 3). Notably we see that the cluster regions are located

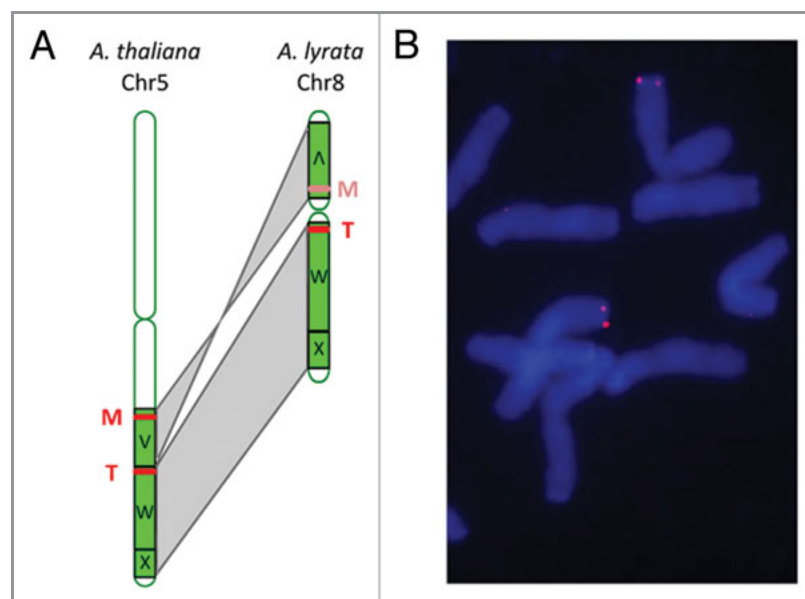


Figure 3. Inferring the ancestral position of the thalianol and marneral gene clusters. (A) The present-day location of the thalianol (T) and marneral (M) clusters in *A. thaliana* and *A. lyrata*, which has an ancestral karyotype. Chromosome maps were prepared using the SynMap tool in CoGe,³⁸ and letters V,W and X refer to the syntenic genomic blocks that make up the genomes of different members of the Brassicaceae.³² (B) DNA FISH of *Avena strigosa* (diploid oat) chromosomes using the gene for the first committed step in avenacin biosynthesis, *Sad1*, as a probe. Image adapted from reference 14.

very near to the chromosome breakpoints that occurred during these rearrangements. Little is known about the nature of chromosome breakpoints in plants. However, in mammals breakpoints are strongly associated with the presence of SDs.³⁰ The cluster region breakpoints also appear to have been repeatedly used during evolution because the chromosome segments that they delimit are found in different chromosomal positions in many different Brassicacea species.³² These observations reinforce our hypothesis that the marneral and thalianol gene clusters lie in dynamic chromosomal regions.

Relative to other members of the Brassicacea *A. thaliana* has an atypical genome that has been reduced in size and chromosome number. Therefore, it is possible that the marneral and thalianol clusters were assembled at a different chromosomal location. Chromosome painting experiments and the sequencing of different Brassicacea genomes together indicate that *A. lyrata* has retained an essentially ancestral karyotype.^{22-24,32} As a consequence, the close proximity of the marneral and thalianol clusters to the centromere of *A. lyrata* chromosome 8 (Fig. 3) is likely to have been preserved in the common ancestor of *A. thaliana* and *A. lyrata* where the gene clusters were first assembled (Fig. 2). If this hypothesis is correct then the centromere border would be a distinctive new location for assembly of a functional gene cluster in plants. Interestingly, while many fungal gene clusters are subtelomeric, in yeast (*Saccharomyces cerevisiae*) the well known galactose utilization pathway (GAL) cluster is located next to the centromere of chromosome 2.³³ We previously noted that there is at least one other putative functional cluster region in *A. thaliana* that contains the *PEN1/BARS1* OSC genes. The *PEN1/BARS1* region is also TE rich, and is located near to the centromere of chromosome 6 in *A. lyrata*. However, the proximity of the *PEN1/BARS1* region to the centromere (approximately 8% of the total chromosome length from the centromere) is not as extreme as we see for the thalianol and marneral clusters (< 1%).

The positioning of the marneral and thalianol clusters points toward a general

mechanism underlying functional gene cluster assembly, because other plant gene clusters are similarly positioned in dynamic chromosomal regions. Three other plant functional gene clusters—the maize cluster for cyclic hydroxamic acids, the sorghum cluster for cyanogenic glucosides, and the oat avenacin cluster (Fig. 3)—are located in subtelomeres. Subtelomeres are well known hotspots for chromosomal recombination and SD. The enhanced turnover of SDs within such regions is likely to provide a special environment or “evolutionary playground” that will accelerate the sampling of different gene combinations, and thus the assembly and functional optimisation of operon-like gene clusters in response to selection pressure.

Of the remaining plant gene clusters three—the *L. japonicus* cyanogenic glucoside cluster and the rice phytocassane and momilactone clusters—are not located in sub-telomeric regions. The location of the cyanogenic glucoside cluster in cassava is not currently known. Further investigations will be needed to find out how these clusters were formed and whether they are also associated with dynamic chromosomal regions.

Better Together?

The primary advantages conferred by clustering of the genes for a metabolic pathway are likely to be co-regulation and co-inheritance (see ref. 14 for a recent review). Little is known about the transcriptional regulation of plant gene clusters. If we take the reasonable assumption that there will be parallels with fungal gene cluster regulation then it is likely that plant gene cluster regulation will be complex and multi-faceted, and will involve cluster-wide chromatin remodelling.³⁴ In relation to the TE-rich Arabidopsis clusters the suspected chromatin-level regulation of the fungal penicillin cluster by a neighboring TE is particularly intriguing.³⁵ Preliminary evidence does support the notion that chromatin-level regulation may be important for the coordination of gene expression in plant functional gene clusters. In oat activation of the avenacin gene cluster is accompanied by chromatin decondensation.³⁶ In *A. thaliana* the

marneral and thalianol cluster genes are marked with histone methylation, and the expression of at least several cluster genes appears to be controlled by chromatin remodelling factors.^{1,3} These findings now need to be built upon and the principal mechanisms behind the transcriptional regulation of these and other plant gene clusters established.

The persistence of secondary metabolite gene clusters in unstable regions is remarkable. Cluster persistence is likely to be promoted by selection for the ability to produce protective compounds. Disruption of these clusters may lead not only to loss of the pathway end products but also to accumulation of toxic/bioactive intermediates, which may further enhance selection for clustering.¹⁴ Furthermore, while subtelomeric, centromere proximal and TE-rich regions may experience relatively high rates of ectopic recombination this does not necessarily extend to increased rates of meiotic recombination. Rather, although the situation in plants is not yet clear, in a number of eukaryotic organisms meiotic recombination is suppressed at centromeres and at many telomeres, and can also be suppressed by repetitive elements that maintain a closed chromatin conformation.³⁷ Therefore, the position and chromatin status of functional gene clusters in plants and in other organisms may serve to protect the cluster regions from breakup during meiosis. Such a feature would have the potential to play a very important role during the fixation of gene clusters in genetically heterogeneous populations.

Conclusions

Over recent years the rate of plant gene cluster discovery has accelerated. The majority of these clusters formed independently, are present in a restricted range of species and can be considered to serve adaptive functions. Our work and the work of others have helped show that similar mechanisms are likely to be behind the formation of the majority of these diverse clusters. We anticipate further important advances into our understanding of gene cluster formation and function as more clusters are analyzed in

detail, and ever more plant genomes are sequenced and scrutinised. The discovery of more clusters may also allow us to

finally establish whether functional gene clusters form only for the production of secondary metabolites in plants or

whether they also form for different pathways and a wider range of adaptive functions.

References

- Field B, Fiston-Lavier AS, Kemen A, Geisler K, Quesneville H, Osbourn AE. Formation of plant metabolic gene clusters within dynamic chromosomal regions. *Proc Natl Acad Sci U S A* 2011; 108:16116-21; PMID:21876149; <http://dx.doi.org/10.1073/pnas.1109273108>
- Hurst LD, Pál C, Lercher MJ. The evolutionary dynamics of eukaryotic gene order. *Nat Rev Genet* 2004; 5:299-310; PMID:15131653; <http://dx.doi.org/10.1038/nrg1319>
- Field B, Osbourn AE. Metabolic diversification-independent assembly of operon-like gene clusters in different plants. *Science* 2008; 320:543-7; PMID:18356490; <http://dx.doi.org/10.1126/science.1154990>
- Horton R, Wilming L, Rand V, Lovering RC, Bruford EA, Khodiyar VK, et al. Gene map of the extended human MHC. *Nat Rev Genet* 2004; 5:889-99; PMID:15573121; <http://dx.doi.org/10.1038/nrg1489>
- Hittinger CT, Rokas A, Carroll SB. Parallel inactivation of multiple GAL pathway genes and ecological diversification in yeasts. *Proc Natl Acad Sci U S A* 2004; 101:14144-9; PMID:15381776; <http://dx.doi.org/10.1073/pnas.0404319101>
- Wong S, Wolfe KH. Birth of a metabolic gene cluster in yeast by adaptive gene relocation. *Nat Genet* 2005; 37:777-82; PMID:15951822; <http://dx.doi.org/10.1038/ng1584>
- Hall C, Dietrich FS. The reacquisition of biotin prototrophy in *Saccharomyces cerevisiae* involved horizontal gene transfer, gene duplication and gene clustering. *Genetics* 2007; 177:2293-307; PMID:18073433; <http://dx.doi.org/10.1534/genetics.107.074963>
- Hoffmeister D, Keller NP. Natural products of filamentous fungi: enzymes, genes, and their regulation. *Nat Prod Rep* 2007; 24:393-416; PMID:17390002; <http://dx.doi.org/10.1039/b603084j>
- Turgeon BG, Bushley KE. Secondary metabolism. In: Borkovich K, Ebbold D, eds. *Cellular and Molecular Biology of Filamentous Fungi*. Washington: American Society of Microbiology, 2010:376-95.
- Frey M, Chomet P, Glawischnig E, Stettner C, Grün S, Winklmair A, et al. Analysis of a chemical plant defense mechanism in grasses. *Science* 1997; 277:696-9; PMID:9235894; <http://dx.doi.org/10.1126/science.277.5326.696>
- Qi X, Bakht S, Leggett M, Maxwell C, Melton R, Osbourn A. A gene cluster for secondary metabolism in oat: implications for the evolution of metabolic diversity in plants. *Proc Natl Acad Sci U S A* 2004; 101:8233-8; PMID:15148404; <http://dx.doi.org/10.1073/pnas.0401301101>
- Qi X, Bakht S, Qin B, Leggett M, Hemmings A, Mellon F, et al. A different function for a member of an ancient and highly conserved cytochrome P450 family: from essential sterols to plant defense. *Proc Natl Acad Sci U S A* 2006; 103:18848-53; PMID:17124172; <http://dx.doi.org/10.1073/pnas.0607849103>
- Mugford ST, Qi X, Bakht S, Hill L, Wegel E, Hughes RK, et al. A serine carboxypeptidase-like acyltransferase is required for synthesis of antimicrobial compounds and disease resistance in oats. *Plant Cell* 2009; 21:2473-84; PMID:19684243; <http://dx.doi.org/10.1105/tpc.109.065870>
- Chu HY, Wegel E, Osbourn A. From hormones to secondary metabolism: the emergence of metabolic gene clusters in plants. *Plant J* 2011; 66:66-79; PMID:21443624; <http://dx.doi.org/10.1111/j.1365-313X.2011.04503.x>
- Wilderman PR, Xu M, Jin Y, Coates RM, Peters RJ. Identification of syn-pimara-7,15-diene synthase reveals functional clustering of terpene synthases involved in rice phytoalexin/allelochemical biosynthesis. *Plant Physiol* 2004; 135:2098-105; PMID:15299118; <http://dx.doi.org/10.1104/pp.104.045971>
- Shimura K, Okada A, Okada K, Jikumaru Y, Ko KW, Toyomasu T, et al. Identification of a biosynthetic gene cluster in rice for momilactones. *J Biol Chem* 2007; 282:34013-8; PMID:17872948; <http://dx.doi.org/10.1074/jbc.M703344200>
- Swaminathan S, Morrone D, Wang Q, Fulton DB, Peters RJ. CYP76M7 is an ent-cassadiene C11alpha-hydroxylase defining a second multifunctional diterpene biosynthetic gene cluster in rice. *Plant Cell* 2009; 21:3315-25; PMID:19825834; <http://dx.doi.org/10.1105/tpc.108.063677>
- Wang Q, Hillwig ML, Peters RJ. CYP99A3: functional identification of a diterpene oxidase from the momilactone biosynthetic gene cluster in rice. *Plant J* 2011; 65:87-95; PMID:21175892; <http://dx.doi.org/10.1111/j.1365-313X.2010.04408.x>
- Takos AM, Knudsen C, Lai D, Kannangara R, Mikkelsen L, Motawia MS, et al. Genomic clustering of cyanogenic glucoside biosynthetic genes aids their identification in *Lotus japonicus* and suggests the repeated evolution of this chemical defence pathway. *Plant J* 2011; 68:273-86; PMID:21707799; <http://dx.doi.org/10.1111/j.1365-313X.2011.04685.x>
- de Silva K, Laska B, Brown C, Sederoff HW, Khodakovskaya M. *Arabidopsis thaliana* calcium-dependent lipid-binding protein (AtCLB): a novel repressor of abiotic stress response. *J Exp Bot* 2011; 62:2679-89; PMID:21252258; <http://dx.doi.org/10.1093/jxb/erq468>
- Beilstein MA, Nagalingum NS, Clements MD, Manchester SR, Mathews S. Dated molecular phylogenies indicate a Miocene origin for *Arabidopsis thaliana*. *Proc Natl Acad Sci U S A* 2010; 107:18724-8; PMID:20921408; <http://dx.doi.org/10.1073/pnas.0909766107>
- Wang X, Wang H, Wang J, Sun R, Wu J, Liu S, et al. Brassica rapa Genome Sequencing Project Consortium. The genome of the mesopolyploid crop species *Brassica rapa*. *Nat Genet* 2011; 43:1035-9; PMID:21873998; <http://dx.doi.org/10.1038/ng.919>
- Dassanayake M, Oh DH, Haas JS, Hernandez A, Hong H, Ali S, et al. The genome of the extremophile crucifer *Thellungiella parvula*. *Nat Genet* 2011; 43:913-8; PMID:21822265; <http://dx.doi.org/10.1038/ng.889>
- Thellungiella halophila* Genome Project. 2011.
- Bowers JE, Chapman BA, Rong J, Paterson AH. Unravelling angiosperm genome evolution by phylogenetic analysis of chromosomal duplication events. *Nature* 2003; 422:433-8; PMID:12660784; <http://dx.doi.org/10.1038/nature01521>
- Freeling M, Lyons E, Pedersen B, Alam M, Ming R, Lisch D. Many or most genes in Arabidopsis transposed after the origin of the order Brassicales. *Genome Res* 2008; 18:1924-37; PMID:18836034; <http://dx.doi.org/10.1101/gr.081026.108>
- Dooner HK, Weil CF. Give-and-take: interactions between DNA transposons and their host plant genomes. *Curr Opin Genet Dev* 2007; 17:486-92; PMID:17919898; <http://dx.doi.org/10.1016/j.cdev.2007.08.010>
- Fiston-Lavier AS, Anxolabehere D, Quesneville H. A model of segmental duplication formation in *Drosophila melanogaster*. *Genome Res* 2007; 17:1458-70; PMID:17726166; <http://dx.doi.org/10.1101/gr.6208307>
- Hu TT, Pattyn P, Bakker EG, Cao J, Cheng JF, Clark RM, et al. The *Arabidopsis lyrata* genome sequence and the basis of rapid genome size change. *Nat Genet* 2011; 43:476-81; PMID:21478890; <http://dx.doi.org/10.1038/ng.807>
- Marques-Bonet T, Giritrajan S, Eichler EE. The origins and impact of primate segmental duplications. *Trends Genet* 2009; 25:443-54; PMID:19796838; <http://dx.doi.org/10.1016/j.tig.2009.08.002>
- Amoutzias G, Van de Peer Y. Together we stand: genes cluster to coordinate regulation. *Dev Cell* 2008; 14:640-2; PMID:18477446; <http://dx.doi.org/10.1016/j.devcel.2008.04.006>
- Lysak MA, Koch MA. Phylogeny, genome, and karyotype evolution of crucifers (Brassicaceae). In: Schmidt R, Bancroft I, eds. *Genetics and Genomics of the Brassicaceae*. New York: Springer, 2011:1-31.
- Bassel J, Mortimer R. Genetic order of the galactose structural genes in *Saccharomyces cerevisiae*. *J Bacteriol* 1971; 108:179-83; PMID:5122803
- Yin W, Keller NP. Transcriptional regulatory elements in fungal secondary metabolism. *J Microbiol* 2011; 49:329-39; PMID:21717315; <http://dx.doi.org/10.1007/s12275-011-1009-1>
- Shaaban M, Palmer JM, El-Naggar WA, El-Sokkary MA, Habib SE, Keller NP. Involvement of transposon-like elements in penicillin gene cluster regulation. *Fungal Genet Biol* 2010; 47:423-32; PMID:20219692; <http://dx.doi.org/10.1016/j.fgb.2010.02.006>
- Wegel E, Koumproglou R, Shaw P, Osbourn A. Cell type-specific chromatin decondensation of a metabolic gene cluster in oats. *Plant Cell* 2009; 21:3926-36; PMID:20040536; <http://dx.doi.org/10.1105/tpc.109.072124>
- Lichten M, de Massy B. The impressionistic landscape of meiotic recombination. *Cell* 2011; 147:267-70; PMID:22000007; <http://dx.doi.org/10.1016/j.cell.2011.09.038>
- Lyons E, Pedersen B, Kane J, Alam M, Ming R, Tang H, et al. Finding and comparing syntenic regions among Arabidopsis and the outgroups papaya, poplar, and grape: CoGe with rosids. *Plant Physiol* 2008; 148:1772-81; PMID:18952863; <http://dx.doi.org/10.1104/pp.108.124867>