Legume Genomics and Genetics 2024, Vol.15, No.1, 37-44
http://cropscipublisher.com/index.php/lgg

Feature Review **Open Access**

Impact of Whole Genome Duplication Events on the Diversification of Legumes

Weiliang Shen, Yupin Huang, Ruchuan Chen, Hangming Lin
Tropical Legume Research Center, Hainan Institute of Tropical Agricultural Resources, Sanya, 572025, Hainan, China
Corresponding email: hangming.lin@hitar.org

Legume Genomics and Genetics, 2024 Vol.15, No.1 doi: 10.5376/lgg.2024.15.0005
Received: 12 Jan., 2024
Accepted: 13 Feb., 2024
Published: 24 Feb., 2024

Preferred citation for this article:
Shen W.L., Huang Y.P., Chen R.C., and Lin H.M., 2024, Impact of whole genome duplication events on the diversification of legumes, Legume Genomics and Genetics, 15(1): 37-44 (doi: 10.5376/lgg.2024.15.0005)

Abstract This study aims to explore the role of whole genome duplication (WGD) events in the evolutionary history and diversification of legumes. It seeks to summarize the mechanisms, historical occurrences, and impacts of WGD on genetic diversity, ecological adaptation, and agricultural significance in legumes. The study identifies key WGD events in the evolutionary timeline of legumes and discusses their mechanisms, including autopolyploidy and allopolyploidy. It highlights the significant genetic and evolutionary consequences of WGD, such as enhanced genetic variation, novel trait development, and increased adaptability to diverse environments. Additionally, it examines the impact of WGD on legume diversification at both the ecological and functional levels, noting specific examples within major legume subfamilies. Whole genome duplication events have played a crucial role in shaping the evolutionary trajectory and diversification of legumes. These events have contributed to genetic innovation, ecological niche expansion, and the development of economically important traits. The study emphasizes the importance of further research to fully understand the functional implications of WGD and its potential applications in legume breeding and conservation.

Keywords Whole genome duplication (WGD); Legume diversification; Genetic diversity; Evolutionary adaptation; Agricultural significance

1 Introduction

The legume family, Fabaceae, is the third-largest family of angiosperms and holds significant economic and ecological importance. Legumes are the second-most-important crop family, providing essential nutrients and serving as a key source of biological nitrogen in agriculture (Young and Bharti, 2012; Vlasova et al., 2016). The family includes agriculturally vital species such as soybeans, common beans, and peanuts, which are crucial for global food security and sustainable agricultural practices. The ability of legumes to fix atmospheric nitrogen through symbiotic relationships with rhizobia bacteria enhances soil fertility, reducing the need for synthetic fertilizers and promoting environmentally friendly farming (Young and Bharti, 2012; Cannon et al., 2015).

Whole genome duplication (WGD) refers to events where an organism's entire genome is duplicated, resulting in polyploidy. These events can be categorized into autopolyploidy, where the duplication occurs within a single species, and allopolyploidy, where it involves hybridization between different species. WGDs have played a pivotal role in the evolution of plants, contributing to genetic diversity, speciation, and the development of novel traits (Young and Bharti, 2012; Cannon et al., 2015). In legumes, multiple WGD events have been identified, including a significant event approximately 58 million years ago that is shared by most agriculturally important legumes (Young and Bharti, 2012; Vlasova et al., 2016). These duplications have led to gene fractionation, subfunctionalization, and neofunctionalization, which have been instrumental in shaping legume genomes and their unique characteristics (Young and Bharti, 2012; Vlasova et al., 2016).

This study is to explore the impact of whole genome duplication events on the diversification of legumes. This study synthesize current knowledge on the timing, occurrence, and evolutionary consequences of WGDs in legumes, with a particular focus on their role in the development of legume-specific traits and their ecological and agricultural significance. Understanding the influence of WGDs on legume diversification is crucial for advancing our knowledge of plant evolution and for improving legume breeding strategies to enhance crop resilience and productivity.

2 Whole Genome Duplication

2.1 Mechanisms of WGD

Whole genome duplication (WGD) is a significant evolutionary event that results in the doubling of an organism's entire genome. This process can occur through two primary mechanisms: autopolyploidy and allopolyploidy. Autopolyploidy involves the duplication of a genome within a single species, leading to multiple sets of chromosomes that are identical. In contrast, allopolyploidy results from the hybridization between two different species, followed by chromosome doubling, which combines the genomes of both parent species (McGrath and Lynch, 2012; Ren et al., 2018).

The molecular and genetic basis of WGD involves complex processes that include errors in meiosis or mitosis, leading to the formation of polyploid cells. These errors can be due to the failure of chromosome segregation or cytokinesis, resulting in cells with double the normal chromosome number. The retention of duplicated genes following WGD is influenced by several factors, including dosage balance constraints, which favor the maintenance of gene duplicates to preserve the balance of gene products (McGrath and Lynch, 2012).

2.2 Evolutionary consequences of WGD

The evolutionary consequences of WGD are profound and multifaceted. One of the primary outcomes is gene duplication, which provides raw material for evolutionary innovation. Duplicated genes can undergo various fates, including pseudogenization (loss of function), neofunctionalization (acquisition of new functions), and subfunctionalization (partitioning of the original function between duplicates) (McGrath and Lynch, 2012; Ren et al., 2018). These processes contribute to functional diversification and can lead to the evolution of new traits and adaptations.

WGD also impacts genome structure and stability. The presence of duplicated genes can lead to genomic rearrangements and changes in chromosome structure, which may affect genome stability. However, the retention of duplicated genes can also enhance genome robustness by providing redundancy and buffering against deleterious mutations (McGrath and Lynch, 2012).

Furthermore, WGD plays a crucial role in speciation and adaptation. The loss of duplicate genes following WGD can lead to reproductive isolation between populations, thereby increasing the speciation rate. This process is particularly significant in polyploid lineages, where WGD can drive rapid diversification and adaptation to new environments (McGrath and Lynch, 2012; Ren et al., 2018). The preferential retention of genes involved in essential cellular functions and metabolic pathways following WGD suggests that these events can enhance the adaptive potential of species, allowing them to thrive in changing environmental conditions (Ren et al., 2018).

In summary, WGD is a pivotal evolutionary mechanism that drives gene duplication, functional diversification, and species diversification. Its impact on genome structure, stability, and adaptation underscores its significance in the evolutionary history of legumes and other angiosperms.

3 Historical WGD Events in Legumes

3.1 Identification of WGD events in legumes

Whole genome duplication (WGD) events in legumes have been identified using a variety of genomic and phylogenomic methods. Comparative genomic analyses, which involve comparing the genomes of different species to identify duplicated regions, are commonly used. Phylogenomic analyses, which construct phylogenetic trees from gene family data, help in pinpointing the timing and occurrence of WGDs. For instance, one study utilized comparative genomic and phylogenomic analyses of 59 public genomes/transcriptomes and 46 newly sequenced transcriptomes to detect large-scale gene duplication events in angiosperms, including legumes (Figure 1) (Ren et al., 2018). Another study employed a combination of phylogenetic methods, such as the stepwise AIC approach (MEDUSA), Bayesian mixture model approach (BAMM), and state-dependent diversification analyses (MuSSE), to investigate patterns of diversification following WGD events (Landis et al., 2018).

Several key WGD events have been identified in the evolutionary history of legumes. One significant WGD event occurred in the common ancestor of the Papilionoideae subfamily, which includes important crops like soybean, peanut, and common bean (Cannon et al., 2015). This WGD event is believed to have contributed to the diversification and success of this subfamily. Additionally, independent WGDs have been detected near the base of other major legume lineages, such as the Mimosoideae-Cassiinae-Caesalpinieae (MCC) clade, Detarieae, and *Cercideae clades* (Cannon et al., 2015). These WGDs are associated with the emergence of novel traits and increased species richness in these lineages.

3.2 Case studies of WGD in specific legume lineages

The Papilionoideae subfamily, also known as the Faboideae, has experienced significant WGD events that have shaped its evolution. The WGD event in the common ancestor of all papilionoids is a notable example. This event has been linked to the diversification of nodulating and non-nodulating taxa within the subfamily, including economically important crops such as alfalfa, soybean, and common bean (Cannon et al., 2015; Nadon and Jackson, 2020). The polyploid origins of these crops have implications for their domestication and improvement, as polyploidy can generate new functional diversity and drive speciation (Nadon and Jackson, 2020).

In the Mimosoideae subfamily, WGDs have also played a crucial role in diversification. The MCC clade, which includes the Mimosoideae, experienced an ancestral WGD event that is associated with the emergence of nodulation, a key trait for nitrogen fixation (Cannon et al., 2015). This WGD event likely contributed to the ecological success and diversification of the Mimosoideae subfamily.

The Caesalpinioideae subfamily has also been influenced by WGD events. Independent WGDs have been detected near the base of the Caesalpinioideae lineage, contributing to the diversification of this group (Cannon et al., 2015). These WGDs are associated with the development of novel traits and increased species richness, similar to the patterns observed in other legume subfamilies.

In summary, WGD events have had a profound impact on the diversification of legumes, with significant events identified in the Papilionoideae, Mimosoideae, and Caesalpinioideae subfamilies. These events have contributed to the emergence of novel traits, increased species richness, and the ecological success of these lineages.

4 Impacts of WGD on Legume Diversification

4.1 Genetic diversity and evolutionary innovation

Whole genome duplication (WGD) events have been pivotal in enhancing genetic diversity and fostering evolutionary innovation in legumes. WGDs result in the doubling of the entire genome, providing raw genetic material that can lead to the development of novel traits and adaptations. This process is crucial for the acquisition of evolutionary novelties, which are essential for adapting to new environments and invading new ecological niches (Moriyama and Koshiba-Takeuchi, 2018). The retention of duplicated genes following WGD can facilitate phenotypic evolution, allowing legumes to develop unique traits that contribute to their diversification (Ren et al., 2018; Landis et al., 2018). For instance, the preferential retention of genes involved in essential cellular metabolisms post-WGD has been linked to the promotion of angiosperm radiation and enhanced adaptation to environmental changes (Ren et al., 2018).

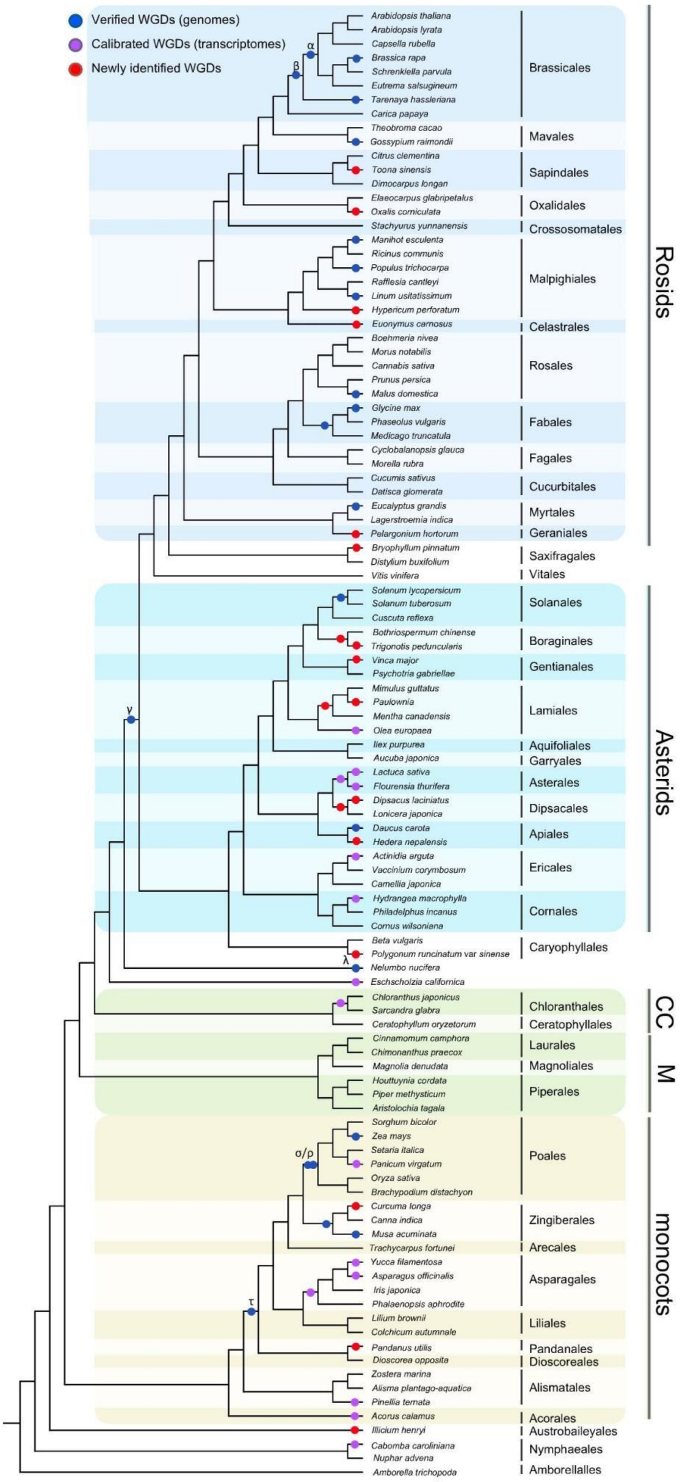

Figure 1 Illustration of phylogenetic relationships of 105 angiosperms and proposed WGDs (Adopted from Ren et al., 2018)

Image caption: The phylogenetic tree depicts the relationships among 105 angiosperm species and highlights proposed whole-genome duplications (WGDs); Legend: Blue circles represent verified WGDs based on genomes; Purple circles denote calibrated WGDs based on transcriptomes; Red circles indicate newly identified WGDs; The tree is divided into major clades, including Rosids, Asterids, and monocots, showing the evolutionary history and the occurrence of WGDs within these groups. This comprehensive illustration provides insight into the phylogenetic relationships and genomic evolution of angiosperms (Adopted from Ren et al., 2018)

4.2 Ecological and functional diversification

WGDs have also played a significant role in the ecological and functional diversification of legumes. The increased genetic variation resulting from WGD allows legumes to adapt to a variety of environmental conditions and develop unique ecological niches. This adaptability is evident in the ability of polyploid lineages to tolerate increased stress and changes in cell size, which can challenge the function of certain cell types but also provide new adaptive advantages (Bomblies et al., 2020). Additionally, WGDs have been associated with shifts in diversification rates, with younger WGD events more likely to be followed by an upshift in diversification, indicating a strong link between WGD and ecological success (Landis et al., 2018). The ability to adapt to different environments and develop unique niches has been a driving force behind the extensive diversification observed in legume species.

4.3 Agricultural and economic implications

The agricultural and economic implications of WGD in legumes are profound. WGDs have contributed to the domestication and improvement of legume crops by enhancing traits such as yield, stress resistance, and nutritional value. The retention of duplicated genes can lead to increased genetic versatility, which is crucial for the development of traits beneficial for agriculture (Hofberger et al., 2013). For example, the diversification of metabolic pathways following WGD can result in improved stress tolerance and better adaptation to agricultural environments (Hoek and Hogeweg, 2009). Furthermore, the immediate fitness advantages conferred by WGD can lead to the rapid development of beneficial traits, although there may be trade-offs with long-term adaptability (Fisher et al., 2018). Overall, WGDs have been instrumental in shaping the genetic architecture of legume crops, making them more resilient and productive, which has significant implications for food security and agricultural sustainability.

In summary, whole genome duplication events have had a profound impact on the diversification of legumes by enhancing genetic diversity, facilitating ecological and functional diversification, and contributing to agricultural and economic improvements. These events have provided the genetic foundation for the development of novel traits and adaptations, enabling legumes to thrive in diverse environments and become valuable agricultural resources.

5 Comparative Genomics and Phylogenetics of WGD in Legumes

5.1 Advances in genomic technologies

Recent advancements in genomic technologies have significantly enhanced our understanding of whole genome duplications (WGDs) in legumes. High-throughput sequencing techniques, such as next-generation sequencing (NGS) and transcriptome sequencing, have enabled the generation of extensive genomic and transcriptomic datasets. These datasets are crucial for detecting and analyzing WGDs across various legume species. For instance, Ren et al. (2018) utilized 59 public genomes/transcriptomes and 46 newly sequenced transcriptomes to identify large-scale gene duplication events in angiosperms, including legumes. This comprehensive approach has confirmed previously reported WGDs and uncovered novel ones, providing a robust framework for further genomic studies.

Bioinformatics tools have also played a pivotal role in WGD analysis. Tools for phylogenomic analyses, such as the Bayesian mixture model approach (BAMM) and state-dependent diversification analyses (MuSSE), have been employed to investigate the patterns of diversification following WGD events. Landis et al. (2018) utilized these methods to analyze the association between WGDs and diversification rates in angiosperms, revealing significant insights into the evolutionary impact of WGDs. These tools facilitate the reconstruction of evolutionary histories and the identification of key genomic changes associated with WGDs.

5.2 Phylogenetic relationships and WGD

Comparative genomics has provided valuable insights into the phylogenetic relationships and evolutionary histories of legumes in the context of WGDs. The study by Cannon et al. (2015) investigated the timing of

WGDs in the legume family by analyzing phylogenetic trees constructed from transcriptome and genome data. This research revealed that the papilionoid WGD occurred in the common ancestor of all papilionoids, highlighting the significance of WGDs in the diversification of this major legume lineage. Additionally, evidence for several independent WGDs near the base of other major legume lineages, such as the Mimosoideae-Cassiinae-Caesalpinieae (MCC) clade, was found, further emphasizing the widespread occurrence of WGDs in legumes.

The reconstruction of evolutionary histories has also been enhanced by the integration of genomic and phylogenetic data. For example, the study by Schranz et al. (2012) proposed the WGD Radiation Lag-Time Model, which suggests that WGDs contribute to the origin of novel key traits and drive species diversification. This model posits that there are significant lag-times between WGDs and subsequent radiations, with the ultimate success of species-rich clades being influenced by later evolutionary phenomena, such as migration events and changing environmental conditions.

In conclusion, advances in genomic technologies and bioinformatics tools have greatly facilitated the study of WGDs in legumes. Comparative genomics and phylogenetic analyses have provided critical insights into the evolutionary impact of WGDs, revealing their role in the diversification and adaptation of legume species. These findings underscore the importance of WGDs in shaping the evolutionary trajectories of legumes and other angiosperms.

6 Challenges and Future Directions in WGD Research

6.1 Current limitations and gaps in knowledge

One of the primary challenges in studying whole genome duplication (WGD) events is the incomplete genomic data available for many species. Despite the increasing number of sequenced genomes, there are still significant gaps, particularly in non-model organisms. This lack of comprehensive genomic data hinders the ability to detect and analyze WGDs accurately across diverse lineages (Landis et al., 2018; Ren et al., 2018). The limited availability of high-quality, annotated genomes restricts our understanding of the full extent and impact of WGDs on species diversification and adaptation (Ren et al., 2018).

Detecting and analyzing WGDs is inherently complex due to the intricate nature of polyploid genomes. WGDs result in the duplication of entire genomes, leading to challenges in distinguishing between ancient and recent duplication events. Additionally, the subsequent gene loss and rearrangement further complicate the identification of WGD events (Schranz et al., 2012; Landis et al., 2018). Advanced computational methods and robust phylogenomic analyses are required to accurately infer WGDs and their evolutionary consequences, but these methods are still evolving and often require extensive computational resources (Smet and Peer, 2012; Landis et al., 2018).

6.2 Future research priorities

To overcome the current limitations, future research should focus on integrating multi-omics approaches, including genomics, transcriptomics, proteomics, and metabolomics. This integrative strategy can provide a more comprehensive understanding of the functional consequences of WGDs by revealing how duplicated genes are expressed, regulated, and interact within cellular networks (Smet and Peer, 2012; Mottes et al., 2021). Utilizing high-throughput data and advanced bioinformatics tools will enable researchers to dissect the complex molecular mechanisms underlying WGD events and their impact on organismal biology (Smet et al., 2012).

Another critical area for future research is exploring the functional consequences of WGDs. While it is known that WGDs can lead to novel functions and increased biological complexity, the specific roles of duplicated genes in adaptive evolution and species diversification remain underexplored (Schranz et al., 2012; Ren et al., 2018). Investigating how WGDs contribute to the development of key traits and ecological adaptations will provide valuable insights into the evolutionary significance of these events (Schranz et al., 2012; Veitia and Birchler, 2021).

Understanding the impact of WGDs on plant and animal genomes has significant implications for breeding and conservation efforts. In agriculture, leveraging knowledge of WGDs can aid in the development of crops with enhanced traits such as increased yield, stress tolerance, and disease resistance (Schranz et al., 2012; Ren et al., 2018). In conservation biology, recognizing the role of WGDs in species diversification can inform strategies to preserve genetic diversity and manage endangered species (Schranz et al., 2012). Future research should aim to translate the findings from WGD studies into practical applications that benefit both agriculture and biodiversity conservation (Schranz et al., 2012; Ren et al., 2018).By addressing these challenges and focusing on these research priorities, the scientific community can advance our understanding of WGDs and their profound impact on the diversification and adaptation of legumes and other plant lineages.

7 Concluding Remarks

Whole genome duplication (WGD) events have been identified as significant drivers of genome complexity and species diversification in legumes. Multiple studies have confirmed the occurrence of WGDs in various legume lineages, including the Papilionoideae and Mimosoideae clades, which have experienced ancestral WGDs . These events have been linked to increased rates of diversification and the emergence of novel traits, which have facilitated the adaptation of legumes to diverse environments. Additionally, WGDs have been shown to result in the retention of duplicate genes that play crucial roles in essential cellular processes and regulatory functions, further contributing to the evolutionary success of legumes .

The significance of WGD in legume evolution is underscored by its role in promoting genetic diversity and enabling the development of key traits that have allowed legumes to thrive in various ecological niches. WGDs have provided raw genetic material for evolutionary innovation, leading to the acquisition of novel phenotypic traits and enhanced adaptability. The timing of WGDs in relation to major climatic changes and the origin of symbiotic nitrogen fixation suggests that these events have been pivotal in shaping the evolutionary trajectory of legumes. Furthermore, the preferential retention of genes involved in transcriptional regulation and other critical functions highlights the importance of WGDs in maintaining genome stability and facilitating complex regulatory networks.

Future research on WGDs in legumes should focus on elucidating the precise mechanisms by which these events contribute to species diversification and adaptation. Advances in genomic and transcriptomic technologies will enable more detailed investigations into the functional divergence of duplicate genes and their roles in legume evolution. Additionally, exploring the interplay between WGDs and environmental factors will provide deeper insights into how these events have shaped the evolutionary history of legumes. Continued comparative analyses across diverse legume species and their relatives will further our understanding of the evolutionary significance of WGDs and their impact on the diversification of this important plant family.

Acknowledgments

The authors extend sincere thanks to two anonymous peer reviewers for their invaluable feedback on the manuscript.

Conflict of Interest Disclosure

The authors affirm that this research was conducted without any commercial or financial relationships that could be construed as a potential conflict of interest.

References

Bomblies K., 2020, When everything changes at once: finding a new normal after genome duplication, Proceedings of the Royal Society B: Biological Sciences, 287(1939): 20202154.
https://doi.org/10.1098/rspb.2020.2154

Cannon S., McKain M., Harkess A., Nelson M., Dash S., Deyholos M., Peng Y., Joyce B., Stewart C., Rolf M., Kutchan T., Tan X., Chen C., Zhang Y., Carpenter E., Wong G., Doyle J., and Leebens-Mack J., 2015, Multiple polyploidy events in the early radiation of nodulating and nonnodulating legumes, Molecular Biology and Evolution, 32(1): 193-210.
https://doi.org/10.1093/molbev/msu296

Fisher K., Buskirk S., Vignogna R., Marad D., and Lang G., 2018, Adaptive genome duplication affects patterns of molecular evolution in *Saccharomyces cerevisiae*, PLoS Genetics, 14(5): e1007396.
https://doi.org/10.1371/journal.pgen.1007396

Hoek M., and Hogeweg P., 2009, Metabolic adaptation after whole genome duplication,Molecular Biology and Evolution, 26(11): 2441-2453.
https://doi.org/10.1093/molbev/msp160

Hofberger J., Hofberger J., Lyons E., Edger P., Pires J., and Schranz M., 2013, Whole genome and tandem duplicate retention facilitated glucosinolate pathway diversification in the mustard family, Genome Biology and Evolution, 5(11): 2155-2173.
https://doi.org/10.1093/gbe/evt162

Landis J., Soltis D., Li Z., Marx H., Barker M., Tank D., and Soltis P., 2018, Impact of whole-genome duplication events on diversification rates in angiosperms, American Journal of Botany, 105(3): 348-363.
https://doi.org/10.1002/ajb2.1060

McGrath C., and Lynch M., 2012, Evolutionary significance of whole-genome duplication, In Polyploidy and Genome Evolution, Berlin, Heidelberg: Springer Berlin Heidelberg, pp.1-20.
https://doi.org/10.1007/978-3-642-31442-1_1

Moriyama Y., and Koshiba-Takeuchi K., 2018, Significance of whole-genome duplications on the emergence of evolutionary novelties, Briefings in Functional Genomics, 17: 329-338.
https://doi.org/10.1093/bfgp/ely007

Mottes F., Villa C., Osella M., and Caselle M., 2021, The impact of whole genome duplications on the human gene regulatory networks, PLoS Computational Biology, 17(12): e1009638.
https://doi.org/10.1371/journal.pcbi.1009638

Nadon B., and Jackson S., 2020, The polyploid origins of crop genomes and their implications: a case study in legumes, Advances in Agronomy, 159: 275-313.
https://doi.org/10.1016/BS.AGRON.2019.08.006

Ren R., Wang H., Guo C., Zhang N., Zeng L., Chen Y., Ma H., and Qi J., 2018, Widespread whole genome duplications contribute to genome complexity and species diversity in angiosperms, Molecular Plant, 11(3): 414-428.
https://doi.org/10.1016/j.molp.2018.01.002

Schranz M., Mohammadin S., and Edger P., 2012, Ancient whole genome duplications, novelty and diversification: the WGD Radiation Lag-Time Model, Current opinion in Plant Biology, 15(2): 147-53.
https://doi.org/10.1016/j.pbi.2012.03.011

Smet R., and Peer Y., 2012, Redundancy and rewiring of genetic networks following genome-wide duplication events, Current Opinion in Plant Biology, 15(2): 168-76.
https://doi.org/10.1016/j.pbi.2012.01.003

Veitia R., and Birchler J., 2021, Gene-dosage issues: a recurrent theme in whole genome duplication events, Trends in Genetics, 38(1): 1-3.
https://doi.org/10.1016/j.tig.2021.06.006

Vlasova A., Capella-Gutiérrez S., Rendón-Anaya M., Hernandez-Oñate M., Minoche A., Erb I., Camara F., Prieto-Barja P., Corvelo A., Sanseverino W., Westergaard G., Dohm J., Pappas G., Saburido-Álvarez S., Kedra D., González I., Cozzuto L., Gómez-Garrido J., Aguilar-Morón M., Andreu N., Aguilar O., Garcia-Mas J., Zehnsdorf M., Vázquez M., Delgado-Salinas A., Delaye L., Lowy E., Mentaberry A., Vianello-Brondani R., García J., Alioto T., Sánchez F., Himmelbauer H., Santalla M., Notredame C., Gabaldón T., Herrera-Estrella A., and Guigó R., 2016, Genome and transcriptome analysis of the Mesoamerican common bean and the role of gene duplications in establishing tissue and temporal specialization of genes, Genome Biology, 17: 1-18.
https://doi.org/10.1186/s13059-016-0883-6

Young N., and Bharti A., 2012, Genome-enabled insights into legume biology, Annual Review of Plant Biology, 63: 283-305.
https://doi.org/10.1146/annurev-arplant-042110-103754