Genomic Convergence in the Adaptation to Extreme Environments

Shaohua Xu, Jiayan Wang, Zixiao Guo, Ziwen He* and Suhua Shi*

State Key Laboratory of Biocontrol, Guangdong Provincial Key Laboratory of Plant Resources, Key Laboratory of Biodiversity Dynamics and Conservation of Guangdong Higher Education Institutes, School of Life Sciences, Sun Yat-sen University, Guangzhou, Guangdong, China

*Correspondence: Ziwen He (heziwen@mail.sysu.edu.cn), Suhua Shi (lssssh@mail.sysu.edu.cn)

https://doi.org/10.1016/j.xplc.2020.100117

ABSTRACT

Convergent evolution is especially common in plants that have independently adapted to the same extreme environments (i.e., extremophile plants). The recent burst of omics data has alleviated many limitations that have hampered molecular convergence studies of non-model extremophile plants. In this review, we summarize cases of genomic convergence in these taxa to examine the extent and type of genomic convergence during the process of adaptation to extreme environments. Despite being well studied by candidate gene approaches, convergent evolution at individual sites is rare and often has a high false-positive rate when assessed in whole genomes. By contrast, genomic convergence at higher genetic levels has been detected during adaptation to the same extreme environments. Examples include the convergence of biological pathways and changes in gene expression, gene copy number, amino acid usage, and GC content. Higher convergence levels play important roles in the adaptive evolution of extremophiles and may be more frequent and involve more genes. In several cases, multiple types of convergence events have been found to co-occur. However, empirical and theoretical studies of this higher level convergent evolution are still limited. In conclusion, both the development of powerful approaches and the detection of convergence at various genetic levels are needed to further reveal the genetic mechanisms of plant adaptation to extreme environments.

Key words: convergent evolution, plant genomes, adaptive evolution, extreme environments

Xu S., Wang J., Guo Z., He Z., and Shi S. (2020). Genomic Convergence in the Adaptation to Extreme Environments. Plant Comm. **1**, 100117.

INTRODUCTION

The study of adaptation plays a central role in evolutionary biology. Although adaptive events are often challenging to identify, independent evolution of similar traits in multiple clades is one of the clearest pieces of evidence that adaptation has occurred (Losos, 2011; Martin and Orgogozo, 2013; Stern, 2013; Storz, 2016). Such convergent evolution is common across many taxa. Classic examples include the independent evolution of flight in insects, birds, and bats; eyes in octopi, vertebrates, and spiders; and C_4 photosynthesis in diverse plant lineages. A central question is whether the convergence in phenotype reflects similar underlying molecular events. Most empirical and theoretical work to address this question has focused on unicellular organisms, mammals, model plants, and crops (Zhang, 2006; Li et al., 2010; Lin et al., 2012; Tenaillon et al., 2012; Zhen et al., 2012; Lenser and Theißen, 2013; Stern, 2013; Storz, 2016; Woodhouse and Hufford, 2019; Wu et al., 2020a, 2020b; Zhang et al., 2020). Recently, the availability of genomic data has enabled the expansion of molecular convergence analyses to non-model organisms. The detection

of convergence at the whole-genome level is emerging as a new frontier in evolutionary biology (Parker et al., 2013; Foote et al., 2015; He et al., 2020a; Zhang et al., 2020). However, it is still unknown whether genomic convergence is a common feature of adaptive evolution. Further work is needed to investigate the organization and extent of convergence across multiple genomes to determine its evolutionary frequency.

Plants that have adapted to extreme environments, i.e., extremophiles, are particularly promising subjects for genomic convergent evolution research. Plants are found in virtually all terrestrial habitats, including the most challenging environments. These organisms have established themselves in unstable intertidal zones, dry and hot deserts, aquatic environments, extremely cold arctic environments, and high altitudes with both cold temperatures and strong UV radiation. Stressors from these extreme

environments would be detrimental to most plants, hence there are high selective pressures for traits that allow survival. Having experienced similar strong selective pressures from similar extreme environments, phylogenetically distinct plant taxa have convergently evolved similar traits and adaptive strategies (Hoekstra et al., 2001; Shi et al., 2005; Zhou et al., 2016; Thorogood et al., 2018). With the increasing availability of plant genome data, exploration of the molecular bases that underlie these similar traits has been attempted in several extremophiles (Oh et al., 2012; Wu et al., 2012; Xu et al., 2017a, 2017b; He et al., 2020a).

Despite some limitations, the study of molecular convergence using genomic data not only helps to reveal the molecular basis of phenotypic convergence in extremophiles but also provides insight into the principles of convergent evolution. For example, although adaptive molecular convergence has been detected in a small number of candidate genes, its detection at the whole-genome level has been controversial because of the lack of suitable controls and a large amount of neutral convergence (Parker et al., 2013; Zou and Zhang, 2015a; Foote et al., 2015; Thomas and Hahn, 2015). Extremophiles that have adapted to the same extreme environment and are subject to similar selective pressures may offer greater opportunity for the detection of genomic-level convergent evolution (He et al., 2020b).

Comparison of extremophiles can also help to identify molecular convergence at various genetic levels (Figure 1). The best-studied molecular events are the convergent site substitutions (Zhang and Nei, 1997; Li et al., 2010; Liu et al., 2010; Zhen et al., 2012; Parker et al., 2013; Foote et al., 2015; Hu et al., 2017). Many studies have also found that extremophile plants have undergone the same amino acid (AA) substitutions during adaptation to the same environment (Christin et al., 2010; Xu et al., 2017a; Fukushima et al., 2017). Convergent evolution at other genetic levels has also been detected in extremophiles. For example, an important feature of plant genomes is the high number of gene duplications (Zhang, 2003; Oh et al., 2012). Duplicated genes may evolve new functions or increase the abundance of proteins, both of which have the potential to convergently promote extreme environmental adaptation (Conn et al., 2015; Van Buren et al., 2019). The availability of whole-genome data provides opportunities to comprehensively survey convergent evolution at other genetic levels, such as gene expression variation, TE (transposable element) content, and other features of plant genomes (Yang et al., 2017; Lyu et al., 2018; He et al., 2020a).

In this review, we summarize advances in the study of plant genome convergence during adaptation to extreme environments and explore the extent and the genetic levels of genomic convergence (Table 1). We first discuss convergent site substitutions in a small number of candidate genes, in individual whole genomes, and in population genomics. We then examine genomic convergence at higher levels, focusing on changes in gene copy number, gene expression, and whole-genome composition. We discuss the probability of convergence at these genetic levels. We highlight cases where genomic convergence at various genetic levels contributes to plant adaptation to the same extreme environments and highlight the necessity of developing powerful approaches and comprehensive analysis of convergence at various genetic levels to further advance genomic convergence studies.

CONVERGENCE OF SITE SUBSTITUTIONS

Convergent AA Substitutions in Candidate Genes

Early attempts to detect molecular convergence usually focused on phenotypes associated with known candidate genes. One of the most well-studied examples of convergent evolution among extremophiles is the independent development of C_4 and Crassulacean acid metabolism (CAM) photosynthesis from C_3 photosynthesis. These alternative modes of photosynthesis have evolved in diverse plant species more than 60 times as a result of adaptation to drought and high temperatures (Sage et al., 2011; Edwards and Ogburn, 2012; Heyduk et al., 2019). Several genes involved in photosynthesis appear to have experienced convergent site substitutions. The phosphoenolpyruvate carboxylase (*PEPC*) gene, which encodes a key enzyme in the C_4 photosynthetic pathway, has the highest number of convergent AA substitutions. Among eight independent grass C_4 lineages, 21 *PEPC* sites have converged to similar or identical AAs (Christin et al., 2007). In the second most species-rich C_4 family, Cyperaceae, the *PEPC* gene from 104 C_4 species in five phylogenetically independent groups shares 16 parallel AA substitutions driven by positive selection, five of which also appear in grass and eudicot C_4 lineages (Besnard et al., 2009). Rubisco, an enzyme involved in the fixation of CO_2 into organic compounds, is another key enzyme that has undergone convergent evolution in the photosynthetic system. Rubisco from C_4 plants shows greater catalytic efficiency but lower specificity than that of C_3 plants. This is probably an adaptation to higher CO_2 concentrations in C_4 photosynthetic cells. The *rbcL* gene, which encodes the large subunit of Rubisco, has eight codons with adaptive convergent substitutions among 23 independent monocot C_4 lineages (Christin et al., 2008).

Another example of convergent site substitution occurs in carnivorous plants. Most carnivorous plants are found in damp heaths, bogs, swamps, and muddy or sandy shores where nitrogenous material is often scarce or unavailable because of acidic or other unfavorable soil conditions. To survive in these stressful environments, more than 600 plant species have evolved carnivorous traits; e.g., trap morphology and the presence of digestive fluid proteins (Adamec, 1997; Thorogood et al., 2018). By sequencing digestive fluid proteins from four carnivorous species that originated independently from three clades (*Cephalotus follicularis* in Oxalidales, *Sarracenia purpurea* in Ericales, and *Nepenthes alata* and *Drosera adelae* in Caryophyllales), Fukushima et al. (2017) found that carnivorous plants repeatedly co-opted stress-response genes to putatively enable a particular digestive physiology. The repeatedly utilized proteins (GH19 chitinases, purple acid phosphatases, and RNase T2s) have also acquired convergent AA substitutions. As the convergent substitutions tend to be located at exposed positions that constitute the protein-environment interface, the proteins may experience the same selective pressure from the digestive fluid environment (including the presence of insect-derived substrates, high endogenous proteolytic activity, low pH, and microbial invasion or symbiosis) to maintain catalytic activity.
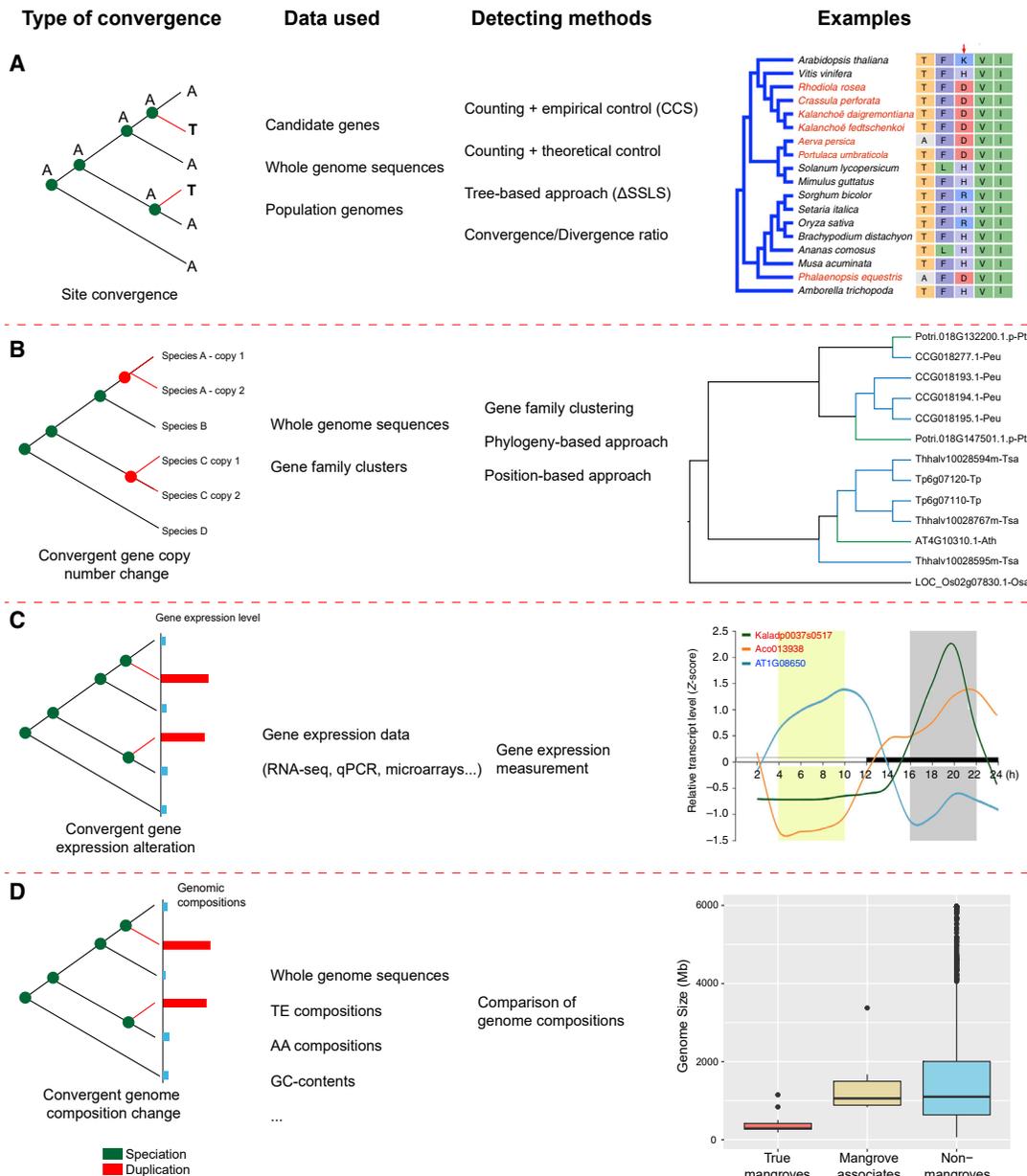
Figure 1. Types of Genomic Convergence.
We classify genomic convergence as events that affect nucleotide substitutions at individual sites, gene copy number, gene expression alteration, and genome composition. The source data, detection methods, and examples are shown for each type. See Box 1 for details of detection methods. The examples in **(A)** and **(C)** are adapted from Yang et al. (2017). The examples in **(B)** and **(D)** are based on data from Ma et al. (2013), Wu et al. (2012), Yang et al. (2013) and Lyu et al. (2018).

The candidate gene approach used to study convergent evolution is usually supported by both evolutionary and functional evidence. Therefore, this approach has identified well-supported convergent events that accompany adaptation to extreme environments. However, the approach is limited to phenotypes associated with pre-existing candidate genes, which are not available for most extremophiles.

Genome-wide Convergence of AA Substitutions
Genomic data provide an opportunity to expand the scope of research to phenotypes with an unknown genetic basis. A series

of approaches developed in candidate genes were applied to detect adaptive genomic convergence (Box 1). In a genome-wide survey of AA substitutions among three CAM plant genomes, Yang et al. (2017) identified four CAM photosynthesis-related genes that had experienced convergent evolution, including the previously identified *PEPC* gene (Figure 1A). There are an unexpectedly large number of convergent AA substitutions in the PEPC proteins of the CAM plants *Kalanchoe fedtschenkoi* and *Phalaenopsis equestris* (Yang et al., 2017). The convergent AA substitutions can increase PEPC activity. Another gene, *HY5* (elongated hypocotyl 5), carries a convergent AA substitution in *K. fedtschenkoi* and

Environmental stress	Plant taxa	Phenotype	Genetic level of convergence				References
			Site substitution	Expression	Copy number	Whole-genome feature	
Drought and hot environment	eight independent grass C$_4$ lineages	C$_4$ photosynthesis	PEPC				Christin et al. (2007)
	C$_4$ lineages in Cyperaceae, grass, and eudicot	C$_4$ photosynthesis	PEPC				Besnard et al. (2009)
	23 independent monocot C$_4$ lineages	C$_4$ photosynthesis	rbcL				Christin et al. (2008)
	A. comosus (pineapple), P. equestris (moth orchid), and K. fedtschenkoi	CAM photosynthesis	PEPC, HY5 (elongated hypocotyl 5) and two CAZyme genes	54 genes including PPCK	both PEPC and PPCK have duplications		Yang et al. (2017)
	C$_4$ lineages in Alloteropsis	C$_4$ photosynthesis			lateral gene transfer of ppc and pck		Christin et al. (2012)
Low nutrient availability in soils	C. follicularis, N. alata, D. adelae, and S. purpurea	carnivorous	GH19 chitinases, purple acid phosphatases, and RNase T2s				Fukushima et al. (2017)
Invasive insect	ash trees (Fraxinus)	EAB (Agrilus planipennis) resistance	53 genes				Kelly et al. (2020)
Highly fluctuating environment of intertidal zone	mangrove taxa: Avicennia marina, Rhizophora apiculata, and Sonneratia alba	High salinity tolerance, vivipary, aerial root	~400 genes			Convergent AA composition change, convergent TE and genome size reduction	Xu et al. (2017a; Lyu et al. (2018); He et al., 2020a
Polar extreme environments	Antarctic psychrophilic green algae (Chlamydomonas sp. ICE-L, and Tetrabaena socialis)	resistance to low temperature	~150 genes				Zhang et al. (2019)
High latitude and low temperature	lodgepole pine (Pinus contorta) and interior spruce (Picea glauca, Picea engelmannii, and their hybrids)	resistance to low temperature	47 genes	61 genes	duplicated genes are enriched of site convergence		Yeaman et al. (2016)
Calamine metalliferous soil	two populations of A. halleri and two populations of A. arenosa	resistance to heavy metal toxicity	24 genes				Arnold et al. (2016); Preite et al. (2019)
High salinity	T salsuginea and P euphratica				tandem duplication of HKT1		Wu et al. (2012); Ma et al. (2013)

Table 1. Summary of Molecular Convergence Events in Plants Adapting to Extreme Environments.

			Genetic level of convergence				
Environmental stress	Plant taxa	Phenotype	Site substitution	Expression	Copy number	Whole-genome feature	References
Desiccation	eight resurrection plants	resurrection			tandem duplication of *ELIPs*		Van Buren et al. (2019)
Parasitic habits	obligate parasitic plants in the Orobanchaceae	host detection	*KAI2d*		*KAI2d* is a rapid-evolving clade of *KAI2* duplication		Conn et al. (2015)
Aquatic environments	loss or simplification of stomata and complex flower structures	*Halophila* and Zosteraceae			convergent loss of ~1200 conservative core genes in monocot		Lee et al. (2018)

Table 1. *Continued*

P. equestris and functions in the blue light signaling pathway, which acts as an input to entrain the circadian clock. Because CAM species temporally separate CO_2 uptake and fixation, the circadian clock should play a key role in controlling the robust oscillation of the physiological and biochemical features of CAM.

Although whole-genome approaches promise to reveal novel convergence events, careful statistical analyses must be performed to distinguish real events from background noise. This problem is more tractable when a small number of candidate genes are analyzed, as background events can be distinguished using neutrality tests. However, whole-genome studies include tens of thousands of genes, only a small fraction of which are expected to acquire adaptive convergent substitutions. The signal from rare events can be overwhelmed by background AA substitutions (Rokas and Carroll, 2008; Goldstein et al., 2015). To robustly control for genome-wide background noise, Xu et al. (2017a) developed a CCS (convergence at conservative sites) method for the detection of genome-wide convergence at only conservative sites (Box 1). The method focuses exclusively on conservative sites, discarding rapidly evolving loci that account for the bulk of background noise. Simulation results strongly suggest that the CCS method has a low false-positive rate. The method can be made even more robust by including controls for each focal taxon to form a symmetrical design. The CCS method has been fruitfully applied to the study of convergence in mangroves, woody plants that originated independently from inland ancestors and have adapted to intertidal environments. Habitats at the interface between terrestrial and marine environments are characterized by high salinity, hypoxia, daily fluctuating tides, strong UV light, high temperature and sedimentation, and muddy anaerobic soils (Giri et al., 2011; He et al., 2019). To colonize new habitats, mangroves have evolved a series of highly specialized characteristics such as salt tolerance, viviparous embryos, and aerial roots (Tomlinson, 1986; Liang et al., 2008). Furthermore, the mangrove suite of traits has evolved many times, with closely related non-mangrove species available as controls. However, despite bringing together the stringent CCS method and a well-suited biological system, almost 80% of the candidate convergent substitutions can be attributed to neutral noise (Xu et al., 2017a). Nevertheless, the study did identify ~400 genes that harbored candidate adaptive convergent AA substitutions (Xu et al., 2017a). The ubiquitin mediated proteolysis, N-glycan biosynthesis, and glutathione metabolism pathways were enriched in convergently evolving genes. These processes are involved in stress response and are likely to be important for mangroves' adaptation to their habitats. He et al., 2020a improved the CCS method to identify genes that harbor convergent AA substitutions. They identified 73 genes that likely underlie salinity tolerance and are located mainly in the plasma membrane (Figure 2A).

The CCS method was also applied to detect convergence in two green algae (i.e., psychrophilic green algae) adapted to extremely cold polar environments (Zhang et al., 2019). Transcriptome sequencing of two independently originated Antarctic psychrophilic green algae (*Chlamydomonas* sp. ICE-L and *Tetrabaena socialis*) identified genes that harbored convergent substitutions. There were 1,031 convergent AA substitutions in psychrophilic algae and 761 in mesophilic algae. The excess convergent substitutions in psychrophilic algae suggest that adaptation to extreme environments is a causal factor in the convergence process. The photosynthetic machinery, multiple antioxidant systems, and several crucial translation elements in Antarctic psychrophilic algae appear to harbor many convergent substitutions, suggesting that these organisms possess relatively stable photosynthetic apparatuses and multiple protective mechanisms. Because of the high degree of neutral convergence, the concentration of candidate convergent genes in whole genomes is a probability statement. For example, even with additional criteria, the probability that any candidate convergent gene identified by Xu et al. (2017a) has experienced convergent evolution is only 50%. Methods innovation and functional tests are needed to explicitly identify true adaptive convergence (He et al., 2020a).

Detection of Site Convergence Using Population Genomic Data

The cases summarized above involve convergent evolution in unrelated or phylogenetically distinct species. Independent adaptation may also occur between closely related species or between

---

Box 1. Methods for Detecting Convergence of AA Substitutions in Whole Genomes.

AA substitutions were the main targets of early attempts to detect convergent substitutions. These studies involve two steps (Zhang and Kumar, 1997). First, AA sites that experienced convergent substitutions are identified. Second, candidate sites are tested for evidence of adaptive evolution against a neutral null hypothesis. Several approaches using different criteria are listed below.

TEST OF WHETHER THE OBSERVED NUMBERS OF CONVERGENT SITES CAN BE ATTRIBUTED TO RANDOM CHANCE

This method estimates the number of expected convergent substitutions under neutral processes using simulations. Computation is performed based on parameters such as the AA substitution rate matrix, AA equilibrium frequencies, and branch lengths (Zou and Zhang, 2015b). Statistical tests are then employed to evaluate whether the number of observed sites can be explained by neutral evolution. When applied to a small number of candidate genes, this method can successfully identify convergent AA substitutions with high confidence (Zhang and Kumar, 1997; Zhang, 2006; Castoe et al., 2009; Li et al., 2010; Zhen et al., 2012; Stern, 2013; Storz, 2016). However, when applied to whole genomes, the expected number of convergent substitutions may vary greatly depending on the AA substitution model selected (Zou and Zhang, 2015b; He et al., 2020b). For example, by ignoring the variation of acceptable AAs among different sites and at different genetic distances, nonadaptive convergence may be underestimated.

Δ SITE-SPECIFIC LIKELIHOOD SUPPORT

The ΔSSLS (site-specific likelihood support) method is a tree-based approach. The method first defines a null hypothesis of a species tree (H_0) and an alternative hypothesis of a convergence tree (H_1). The method then calculates SSLS using maximum likelihood (ML) for both trees. The difference in SSLS between H_0 and H_1 is calculated at each site: ΔSSLS = SSLS (H_0) – SSLS (H_1). The method was first developed to detect the convergence of 13 mitochondrial genes between snakes and agamid lizards (Castoe et al., 2009) and was also applied to detect genome-wide convergence in echolocating mammals by Parker et al. (2013). To calculate genic level ΔSSLS values, Parker et al. (2013) used the mean ΔSSLS of all sites within a gene as the measure of the strength of support for convergent evolution. A negative ΔSSLS value is a signature of convergence. However, the application of the ΔSSLS method to whole-genome data has been shown to produce false-positives because of the lack of suitable controls (Zou and Zhang, 2015a; Thomas and Hahn, 2015).

CONVERGENCE/DIVERGENCE RATIO

A simple and widely used method is the comparison of the ratio of convergence (C) and divergence (D). The C/D ratio test has been widely used as auxiliary evidence of convergence. Under neutral evolution, the number of divergent sites is highly correlated with the number of convergent sites. Within a group of species without convergent evolution, the ratio between each pair of species should be linearly correlated. A pair of species shows evidence of convergent evolution if it has a higher convergent/divergent ratio than other species pairs.
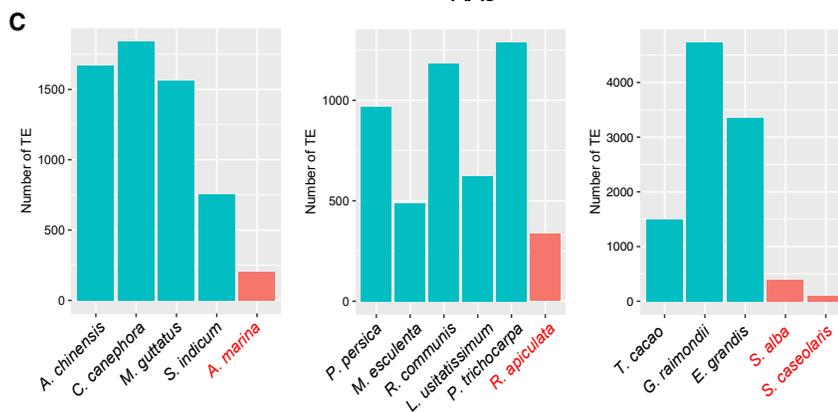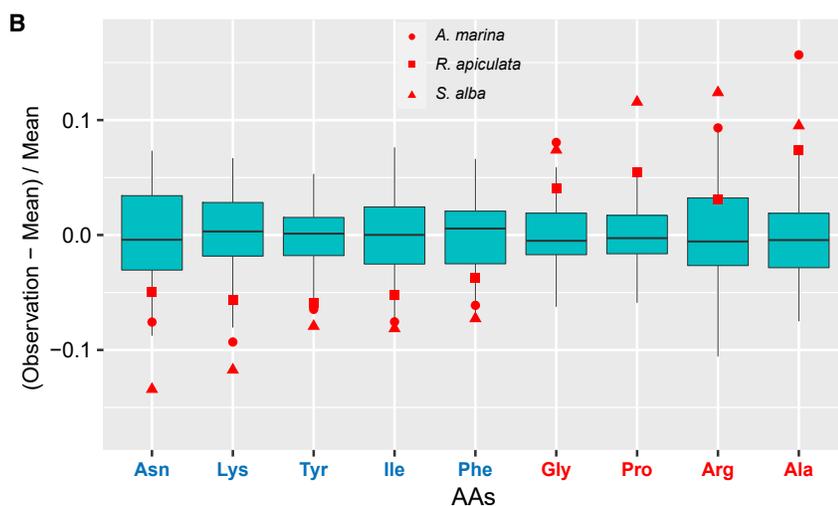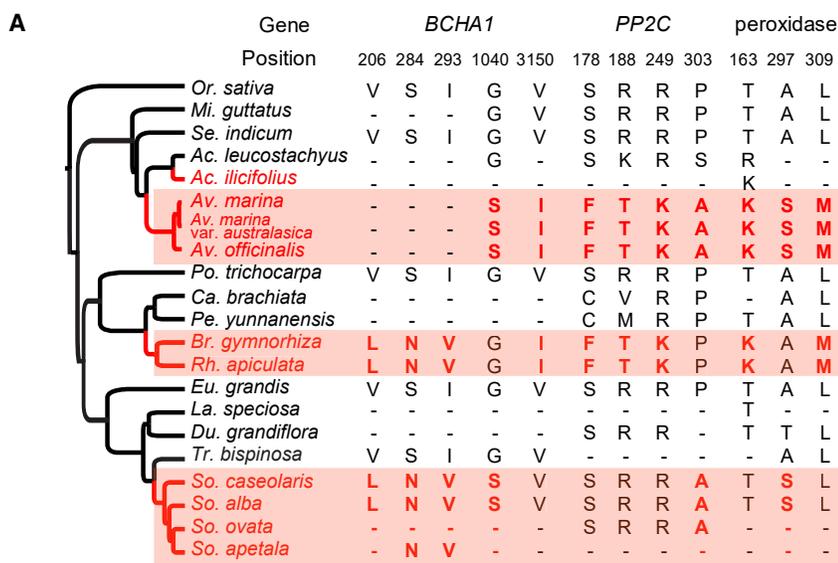
CCS METHOD

The CCS (convergence at conservative sites) method was developed to eliminate the influence of neutral convergence and concentrate on adaptive events (Xu et al., 2017a; He et al., 2020a). The CCS method is based on two principles. First, symmetric phylogenies are used for focal and control groups. The method assumes that convergent evolution is neutral in the control group and can be used as the baseline for detecting adaptive convergence in the focal group. Second, only conservative AA sites are used for convergence detection. The AA sites that are highly variable among species can produce a large number of convergent events by chance. By contrast, conservative sites are usually functionally important, and their convergent substitutions are more likely to be adaptive.

---

populations of the same species. Unlike phylogenetically independent adaptive convergence in which evolution is dominated by independently occurring mutations, the convergent evolution of closely related species or populations can be driven by introgression or by independent selection on standing variation (Stern, 2013). Furthermore, the fitness effects of AA substitutions are often conditional on the genetic background. Thus, convergent site substitutions are more likely to have similar beneficial effects in closely related species (Storz, 2016). The same single nucleotide polymorphisms (SNPs) seem to be repeatedly recruited during adaptation to the same extreme environment, and the comparison of genomic data provides

evidence of both adaptive and independent evolution (Lee and Coop, 2019).

Kelly et al. (2020) exploited the convergent evolution of resistance to the emerald ash borer (EAB, *Agrilus planipennis*) in ash trees (genus *Fraxinus*). By testing for both phenotypic and molecular convergence associated with EAB resistance in *Fraxinus*, they showed that EAB-resistant taxa occur within three independent phylogenetic lineages. They detected 53 genes with evidence of convergent AA evolution in genomes from these resistant lineages. Gene-tree reconstruction indicated that for 48 of these candidates, the convergent AAs are more likely to have arisen by independent evolution than by hybridization or incomplete lineage

	Gene	BCHA1					PP2C				peroxidase		
	Position	206	284	293	1040	3150	178	188	249	303	163	297	309
Or. sativa		V	S	I	G	V	S	R	R	P	T	A	L
Mi. guttatus		-	-	-	G	V	S	R	R	P	T	A	L
Se. indicum		V	S	I	G	V	S	R	R	P	T	A	L
Ac. leucostachyus		-	-	-	G	-	S	K	R	S	R	-	-
Ac. ilicifolius		-	-	-	-	-	-	-	-	-	K	-	-
Av. marina		-	-	-	S	I	F	T	K	A	K	S	M
Av. marina var. australasica		-	-	-	S	I	F	T	K	A	K	S	M
Av. officinalis		-	-	-	S	I	F	T	K	A	K	S	M
Po. trichocarpa		V	S	I	G	V	S	R	R	P	T	A	L
Ca. brachiata		-	-	-	-	-	C	V	R	P	-	A	L
Pe. yunnanensis		-	-	-	-	-	C	M	R	P	T	A	L
Br. gymnorhiza		L	N	V	G	I	F	T	K	P	K	A	M
Rh. apiculata		L	N	V	G	I	F	T	K	P	K	A	M
Eu. grandis		V	S	I	G	V	S	R	R	P	T	A	L
La. speciosa		-	-	-	-	-	-	-	-	-	T	-	-
Du. grandiflora		-	-	-	-	-	S	R	R	-	T	T	L
Tr. bispinosa		V	S	I	G	V	-	-	-	-	-	A	L
So. caseolaris		L	N	V	S	V	S	R	R	A	T	S	L
So. alba		L	N	V	S	V	S	R	R	A	T	S	L
So. ovata		-	-	-	-	-	S	R	R	A	-	-	-
So. apetala		-	N	V	-	-	-	-	-	-	-	-	-

Figure 2. Multi-level Convergence of Mangrove Genomes.

(A) Examples of convergent AA substitutions in three mangrove genomes. Each of the three genes contains at least three convergent AA substitutions and participates in salinity tolerance.

(B) Convergence of AA usage in mangrove genomes. The AA usage of the three mangrove genomes is distinct from that of more than 50 inland dicotyledon genomes. The five most underused AAs are shown on the left (blue font) and the four most overused in mangroves are shown on the right (red font).

(C) The numbers of long terminal repeat-retrotransposons (LTR-RTs) in the three mangrove genomes are convergently smaller than those of their inland relatives. **(A)** and **(B)** are based on data from He et al., 2020a; **(C)** is based on data from Lyu et al. (2018).

temperature. Both the lodgepole pine (*Pinus contorta*) and interior spruce (*Picea glauca*, *Picea engelmannii*, and their hybrids), which diverged approximately 140 million years ago, have evolved a pattern of local adaptation to the climate that reflects a tradeoff between competition for light resources and the acquisition of freezing tolerance. To investigate the basis of this adaptation, Yeaman et al. (2016) searched for correlations between individual SNPs and environmental variables in individuals from >250 populations across the geographic ranges of lodgepole pine and interior spruce. They identified a suite of 47 genes, many of them duplicated, with polymorphisms associated with spatial variation in temperature or cold hardiness in both species. Several of the convergent genes are related to seasonal transitions and abiotic stress (e.g., *PRR5*, *FY*, *FLC*, and *RCAR1*). These results suggest that long-diverged conifers share a suite of genes that play an important role in adaptation to temperature.

In a study of *Arabidopsis* populations that have independently adapted to calamine metalliferous soils, a low degree of convergence was detected in both genes and functional networks (Preite et al., 2019). A high concentration of heavy metal ions is stressful because metalliferous soils are usually nutritionally imbalanced and toxic to plants (Brady et al., 2005; Wójcik et al., 2017). The authors detected convergent genomic footprints of selection in *Arabidopsis halleri* and *Arabidopsis arenosa* at two calamine-type metalliferous (M) sites, each of which was compared with an NM (non-metalliferous) site in its vicinity. These species co-occur at two calamine metalliferous sites with toxic levels of the heavy metals zinc and cadmium. The authors identified five candidate genes with selective sweep signatures that were

sorting. Seven of the candidate genes have putative roles in the phenylpropanoid biosynthesis pathway, which generates products closely related to disease and insect resistance. An additional 17 genes are involved in herbivore recognition, defense signaling, or programmed cell death.

Phylogenetically distant conifers provide an example of genomic convergence as a result of adaptation to high latitude and low

convergent between both population pairs Mias (versus Zapa) and Klet (versus Kowa) in *A. halleri* and another five convergent candidate genes in *A. arenosa*. No convergent candidate genes were common to both species. These results suggest that species-specific metal handling and other biological features may explain the low degree of convergence between species. In another study, the same loci were highly differentiated in both *Arabidopsis lyrata* and *A. arenosa* during adaptation to a serpentine habitat (Arnold et al., 2016).

Artificial domestication is a special kind of extreme environment. Under domestication, strong artificial selection drives diverse domesticated crops to evolve the same traits, such as increased fruit size and seed number, increased yield, loss of seed dormancy, loss of bitterness, and shattering (seed dispersal; Lenser and Theißen, 2013; Woodhouse and Hufford, 2019). The molecular convergence that underlies these shared phenotypic traits has been identified in different plants because of the large quantity of sequencing data available for domesticated plants. For example, the *sh1* gene was found to control the loss of shattering in rice, sorghum, and maize (He et al., 2011; Lin et al., 2012). Because many previous studies on the convergent evolution of domesticated plants have been well summarized (Lenser and Theißen, 2013; Woodhouse and Hufford, 2019), we refer readers to these excellent reviews.

CONVERGENCE OF GENE COPY NUMBER VARIATION

A striking feature of plant genomes is the high proportion of duplicated genes. Gene duplication increases gene expression and can improve tolerance to environmental stressors (Oh et al., 2014; Van De Peer et al., 2017). Furthermore, redundant copies may also accumulate genetic changes and obtain new functions or expression patterns that increase genomic plasticity. For example, the expression of duplicated genes is more variable compared with that of single-copy loci in response to environmental stresses (Ha et al., 2007; Oh et al., 2014). Convergent evolution in both the whole genome and in tandem duplicates has been observed as a result of adaptation to extreme environments (Figure 1B).

Whole-genome duplications (WGDs) occurred not only before the diversification of extant angiosperms and seed plants but also independently in the common ancestors of many groups (Jiao et al., 2011; Wu et al., 2019). By doubling genome content simultaneously, WGDs provide plenty of genetic material for evolution and are proposed to play key roles in adaptation to new environments and diversification (Ohno, 1970; Conant and Wolfe, 2008; Selmecki et al., 2015; Li et al., 2016; Van De Peer et al., 2017; Ren et al., 2018; Wu et al., 2019). Convergent retention of genes related to conditional response to stress and local adaptation may suggest the importance of gene duplication for adaptation to stressful conditions (Li et al., 2016; Van De Peer et al., 2017). A wave of ancient WGD events occurred around the K-Pg (Cretaceous-Paleocene) boundary independently in many plant lineages (Vanneste et al., 2014; Van De Peer et al., 2017; Ren et al., 2018; Wu et al., 2019). Key genes in cold response and shade avoidance pathways convergently maintain duplicated

copies in diverse lineages (Wu et al., 2019). The additional copies evolved new interactions with other genes and promote the survival of plants in cold and dark environments. Modern polyploids, such as tetraploid *Arabidopsis thaliana*, rice, and citrus, have been shown to exhibit increased tolerance to salt or drought stresses (Chao et al., 2013; Yang et al., 2014; Ruiz et al., 2016).

Tandem duplication of individual genes disproportionately affects loci important for stress response (Hanada et al., 2008; Freeling, 2009; Li et al., 2016; Van Buren et al., 2018). Most tandem duplications in plant genomes are relatively new and therefore unlikely to have evolved new functions or expression patterns. They are more likely to benefit survival by increasing absolute transcript abundance. One example is the massive tandem duplication of early light-induced proteins (*ELIPs*) shared by plants tolerant to desiccation (Van Buren et al., 2019). More than 300 diverse resurrection plants with vegetative desiccation tolerance can revive from typically lethal prolonged dehydration (Oliver et al., 2000, 2005). The *ELIP* gene family has expanded convergently through repeated tandem gene duplication in all eight sequenced resurrection plant genomes. Whereas 66 desiccation-sensitive land plants have fewer than 10 *ELIPs*, with an average of 3.1 per genome, eight desiccation-tolerant species have an average of 20.7 *ELIPs* per genome due to independent massive tandem duplications.

The *HKT1* (high-affinity K$^+$ transporter 1) gene has also experienced repeated tandem duplication during adaptation to saline environments (Figure 1B). Compared with only one copy in *A. thaliana*, there are three expressed *HKT1* genes arranged in a tandem gene array in the halophyte *Thellungiella salsuginea* (Wu et al., 2012). A similar tandem duplication has also occurred in the desert poplar, a member of the order Malpighiales that inhabits saline desert environments. The *HKT1* gene family has expanded from one member in the closely related *Populus trichocarpa* genome to four in the *Populus euphratica* genome, with three copies arranged as tandem duplicates (Ma et al., 2013).

Duplicated genes may also acquire new functions (Conant and Wolfe, 2008). Examples include the convergently evolved strigolactone recognition conferred by *D14* and duplicated *KAI2d* genes, which enable host detection in parasitic plants (Conn et al., 2015). Obligate parasitic plants in the Orobanchaceae germinate after sensing plant hormones, strigolactones, exuded from host roots. In *A. thaliana*, the a/b-hydrolase D14 acts as a strigolactone receptor that controls shoot branching, whereas its ancestral paralog KAI2 does not. In a survey of 10 species that represent the full range of parasitism in the Orobanchaceae, more copies of *KAI2* were found in parasite genomes than in those of non-parasitic Lamiales species. The additional copies in parasites formed the rapidly evolving clade *KAI2d*, which had accumulated an excess of AA substitutions. The AA substitutions in *KAI2d* caused protein structure changes and enabled strigolactone recognition.

In addition to gene duplication, other kinds of gene copy number variation also play a role in adaptation to extreme environments. Repeated lateral transfers of genes in the C$_4$ pathway have been

found in the grass lineage *Alloteropsis* (Christin et al., 2012). The vertically acquired C_4 *PEPC* gene in the three C_4 clades is either absent or has no C_4 characteristics. By contrast, the C_4-specific *PEPC* genes from the three C_4 clades each clustered with a distantly related clade. The lateral gene transfer was supported by the high similarity between the intron and untranslated region (UTR) sequences of horizontally acquired genes and those of the putative donors. Another key gene, *pck* (phosphoenolpyruvate carboxykinase), also transitioned from a member of Cenchrinae to the common ancestor of *Alloteropsis angusta* and *Alloteropsis semialata* and now performs a key function in the C_4 pathways of these taxa. The convergent lateral gene transfer may have occurred during direct contact after the deposition of foreign pollen on the *Alloteropsis* stigma. Natural selection may then have led to rapid fixation of the transferred genes in populations.

A decrease in convergent gene copy number (i.e., gene loss) has occurred in seagrasses that have independently returned to the sea. Physiological and morphological features such as stomata and complex floral structures have been lost or simplified in seagrasses due to their re-acquired aquatic life habits (Kuo et al., 2006). The diffusion rate of volatile substances is also reduced underwater. In a survey of two independent seagrass lineages, *Halophila* and Zosteraceae, ~1,200 conserved monocot core genes have been convergently lost. These losses primarily affect genes associated with volatile substance synthesis and signaling and with stomatal development (Lee et al., 2018).

CONVERGENCE OF EXPRESSION REGULATION AND NONCODING VARIATION

During adaptation to stressful environments, molecular phenotypes such as gene expression may have more direct effects on whole-organism traits than site substitutions or gene copy number variations. A large number of plant genes are up- or downregulated in response to environmental stresses (Maruyama et al., 2014; Feng et al., 2020; Gong et al., 2020). However, most of the identified stress-response genes show similar expression changes, and therefore expression regulation is likely to be an ancestral state shared by diverse plant taxa. Convergent gene expression evolution, which is the independent derivation of the same gene expression pattern, is relatively poorly understood because of a lack of relevant studies (Figure 1).

As a key enzyme of C_4 and CAM photosynthesis, *PEPC* not only accumulates extensive convergent AA substitutions but also shows higher expression in all C_4 and CAM plants (Heyduk et al., 2019). Compared with only four genes that show convergent AA substitutions, 54 genes show convergent shifts in diel gene expression patterns in the CAM species *K. fedtschenkoi* and *Ananas comosus* (pineapple; Yang et al., 2017). These significant time shifts in diel transcript expression occur in genes related to nocturnal CO_2 fixation, circadian rhythm, carbohydrate metabolism, stomatal movement, and heat stress response. For example, the maximum transcript abundance of *PPCK* (phosphoenolpyruvate carboxylase kinase 1), a key regulator of *PEPC* that promotes nocturnal CO_2

uptake, shifted from daytime in C_3 photosynthetic species to nighttime in the two CAM species (Figure 1C). The study of convergent evolution of distantly related conifers also revealed convergent gene expression regulation in addition to site substitutions (Yeaman et al., 2016). In response to climate stress, 61 transcripts displayed conserved differential expression patterns in both species. Furthermore, genes with signatures of convergent evolution are also enriched in transcription factors and are more likely to affect the expression of other genes. Convergence of gene expression may also combine with convergence in gene copy number variation. For example, the convergent tandem duplication of *ELIPs* in resurrection plants results in a convergent gene expression pattern change (Van Buren et al., 2019). The expression of *ELIPs is* low or undetectable under well-watered conditions in all surveyed tolerant and sensitive species; however *ELIPs* are consistently among the most highly expressed genes in resurrection plants during dehydration and rehydration. The dramatic expansion of *ELIPs* results in 622-fold higher expression in resurrection plants than in desiccation-sensitive plants under water-deficit conditions. The relatively few examples of gene expression convergence may reflect the scarcity of gene expression data compared with whole-genome sequence data.

CONVERGENCE OF WHOLE-GENOME COMPOSITION

In addition to these functional genes, whole-genome features, such as GC content, AA usage, and TE content, may also be under selective pressure in extreme environments. Using whole-genome sequences from three independent mangrove clades, He et al., 2020a examined the convergence of genomic features related to intertidal environment adaptation. The three mangrove taxa overuse the same set of AAs and underuse others (Figure 2B). Mangroves underuse AAs with large hydrophobic residues whose non-specific inter- and intra-molecular interactions may disrupt proper protein folding and conformation under hypersaline conditions (Paul et al., 2008). The degree of AA usage alteration is highest in the extracellular region where fluctuating salinity has a direct effect. The trend also helps mangroves to reduce energy costs in low-nutrient intertidal environments. The study also found that AA substitution patterns in mangroves are quite different from those of other organisms (He et al., 2020a). Therefore, mangrove proteins have been evolving using a common substitution mechanism that appears to be unique to them, resulting in the observed changes in AA composition.

Lyu et al. (2018) found that all mangrove lineages massively and convergently reduced TE loads, resulting in genome size reduction (Figure 2C). By modeling TE dynamics, they found that the disappearance of these elements is a consequence of fewer births rather than additional deaths. TE-mediated genome instability can accelerate in the face of environmental insults such as high salinity and strong UV radiation (Pfeiffer et al., 2000; Schuermann et al., 2005; Argueso et al., 2008), leading to enhanced long-term selection pressure on the TE load. Therefore, the elimination of most mobile elements leading to a reduced TE load and genome size may be a convergent strategy employed by mangrove during adaptation to new stressful

environments. Furthermore, convergence in transcriptome profiles was found between the two mangroves *Rhizophora mangle* and *Heritiera littoralis* (Dassanayake et al., 2009). Greater GC content has been associated with species that grow in seasonally cold and/or dry climates, perhaps suggesting an advantage of GC-rich DNA during cell freezing and desiccation (Šmarda et al., 2014).

PROBABILITY OF GENOMIC CONVERGENCE IN THE ADAPTATION TO EXTREME ENVIRONMENTS

The numbers of available paths leading to adaptation determine the level of molecular convergence. Relative to the large number of site substitutions across whole genomes, detected convergent substitutions are very rare in plant adaptation to extreme environments. In the two distantly related CAM plant genomes, only five protein-coding genes harbor convergent AA substitutions (Yang et al., 2017). Similar patterns emerge during the adaptation of *Arabidopsis* to calamine metalliferous soil (Preite et al., 2019). In this case, only five convergently evolving genes were detected in two populations that were independently adapting to these habitats, and no convergence was detected between species. The rarity of site convergence may be the result of overly rigid criteria. For example, the criteria used in the CAM plants required that whole genes be clustered into one phylogenetic clade, which is unexpected for genes with a small number of site convergences. On the other hand, site convergences detected at the whole-genome level may be false-positives, given the large degree of neutral convergence. It is still an open question whether the rarity of genomic site convergence is due to a lack of powerful statistical approaches or whether phenotypic convergence involves a small number of underlying DNA sequence changes.

Although convergent evolution is rare at individual sites, it may be more frequent at higher levels of organization. For example, although only four genes contained convergent AA substitutions, 54 genes experienced convergent shifts in diel expression in CAM species (Yang et al., 2017). The higher level of gene expression convergence may reflect the fact that the same expression change can be achieved by various genetic changes, such as gene copy number expansion or mutations in regulatory regions. Convergence in gene copy number may also be more frequent than has been reported among extremophiles. The frequency of gene duplication is much higher in plants than in other domains of life. Gene duplications, especially tandem duplications, can rapidly increase the abundance of their products and may play an important role in short-term adaptation to new environments (Bianconi et al., 2018). Gene duplications produced by WGD have been proposed to promote the adaptation of plants to diverse environments through the evolution of new functions (Van De Peer et al., 2017; Ren et al., 2018; Wu et al., 2019).

In addition to increasing gene copy number, gene duplication may also influence the rate of site convergence, although the relationship between gene duplication and convergent site substitution is still unclear. It is widely believed that convergence results from the minimization of pleiotropy and the maximization of phenotypic change (Stern, 2013). Therefore, gene duplications that produce functional redundancy and alleviate pleiotropic constraints would decrease convergent evolution rates (Zhen et al., 2012; Storz, 2016). This hypothesis is supported by the convergent evolution of resistance to toxic cardenolides through genetic changes in Na^+, K^+-ATPase among diverse insect taxa. In taxa that possess two copies of this gene, a greater number of unique substitutions appear to contribute to cardenolide resistance (Zhen et al., 2012). However, a higher level of site convergence in duplicated genes was observed during adaptation to cold temperatures in conifers. On average, signatures of convergence were 65% more common in cases where one ortholog was duplicated than in one-to-one orthologs (Yeaman et al., 2016). The promotion of adaptation by site convergence in duplicated genes has also been observed in other organisms (Chen et al., 1997; Christin et al., 2007).

Convergent evolution of whole-genome features may affect the majority of protein-coding genes. Selective advantages of AA composition changes may be very small for individual genes but may accumulate to produce large overall effects. Convergent changes in genomic composition, such as AA and GC usage, have been widely explored in microorganisms like halophilic and thermophilic bacteria (Paul et al., 2008). Such convergence may occur in unicellular psychrophilic green algae, whose cells are directly influenced by the environment. Although organs of multicellular plants avoid direct contact with extreme environments, environmental selective pressure may still drive changes in whole-genome composition, such as those observed in mangrove genomes. In addition to the convergence of digestive physiology, convergence of whole-genome composition may also occur in carnivorous plants. For example, nutrient deficiency, which is the major selective pressure on carnivory, may also drive AA composition change that is constrained by nutrition and often selected to achieve cost minimization (Akashi and Gojobori, 2002).

Compared with other convergence types, the convergence of whole-genome composition involves the majority of genes or genomic sequences. Therefore, the convergence of genomic composition requires sufficient time to evolve and is likely to be a relatively long-term adaptive strategy. This may have been the case for the three mangrove taxa discussed above, all of which originated from inland ancestors about 55 million years ago. If they had independently invaded similar environments more recently, perhaps one million and five million years ago, they would not be expected to exhibit comparable convergent signals at the genomic level.

CONCLUSIONS AND FUTURE DIRECTIONS

Adaptive evolution is the central topic of evolutionary biology, and convergent evolution provides the most compelling evidence of adaptive evolution. With the development of sequencing technologies, genomic studies have become prevalent in the field of evolutionary biology, and the large amount of whole-genome data provides opportunities for genomic convergence studies. However, whether such studies can reveal the adaptive mechanisms of genome evolution and to what extent convergence

occurs at the whole genome level are still unknown. In this review, we explored the extent of genomic convergence in plants that have independently adapted to the same extreme environments. Similar and well-defined environmental selection pressures should facilitate the discovery of convergent evolution in these situations. Although convergent evolution studies still focus primarily on model species and experimentation, the accessibility of genomic data has made it possible to test for genomic convergence in non-model species. Indeed, the molecular bases of adaptation in several extremophile plants have been revealed by studying convergent evolution. We summarized the genomic convergence of plants during adaptation to various extreme environmental conditions, such as the unstable intertidal zone, the extreme drought in deserts, aquatic environments, the extreme cold in arctic conditions, pests and pathogens, and cold combined with strong UV at high altitudes.

We found that genomic convergence studies revealed extensive candidate molecular mechanisms of extreme environmental adaptation. Genomic convergence is prevalent at various genetic levels, including site convergence, convergence in gene copy number variation, gene expression, and whole-genome composition. Moreover, this mode of adaptation often co-occurs at several genetic levels. For example, mangrove genomes experienced convergent evolution of AA substitutions, TE and genome size reductions, and AA composition change simultaneously (Xu et al., 2017a; Lyu et al., 2018; He et al., 2020a). CAM plants experienced convergence in both AA substitutions and gene expression alterations (Yang et al., 2017). These studies strongly indicate that genomic convergence is prevalent in extremophile genomes and promises to be a new frontier in comparative genomic studies.

However, there are still challenges in genomic convergence studies. For example, when candidate gene approaches are applied to whole genomes, the large number of neutrally convergent sites can easily overwhelm true convergence. Even when detection is possible, the results are not always reliable because of the lack of suitable controls for neutral convergence (He et al., 2020b). Lateral gene transfer between distantly related species and hybridization between closely related species may also mislead the assessment of convergent evolution (Dunning and Christin, 2020). Therefore, the development of innovative approaches that detect adaptive convergence using genomic data is needed.

More empirical and theoretical studies of convergence at higher genetic levels are also needed. For example, analyses have detected convergence in AA substitution and composition, as well as TE reduction, in several mangrove species from three clades. How much of the adaptation among over 70 mangrove species can be explained by convergent evolution? Do changes in convergent gene copy number and expression pattern also play a role? Similar questions remain in other extremophiles, such as desert plants. In addition to the convergence at these genetic levels, convergence in other genetic changes, such as epigenetic modifications, expression regulation networks, 3D structure of proteins, and sequence insertions and deletions, may also play a role in adaptive convergence and could be explored in future studies. Compared with site convergence, convergence at higher genetic levels lacks a theoretical basis.

Most importantly, the comprehensive analysis of multi-level genomic convergence is still quite limited. In the comprehensive analysis of multi-level genomic convergence, we could examine the order of occurrence of different convergence types during adaptation to a new environment and ask whether large or small effect convergence is more likely to occur. The answers to these questions will help to reveal not only the mechanisms of genomic convergence but also the roles of different genetic variations in adaptive evolution. To analyze multi-level genomic convergence, the collection of multidimensional data is crucial. In addition to the assembly of reference genomes, future studies should also design experiments to generate multidimensional data, such as gene expression and population genetic variations.

In this review, we have focused on plant genome convergence in extreme environments; i.e., under strong selective pressures. Genomic convergence studies could also be applied to a wider range of environmental adaptation processes. Such studies may help to explore the extent of genomic convergence in situations where there is weak selective pressure and determine whether convergence is a common feature of adaptation to similar environments, even when they are not extreme.

Genomic convergence studies not only provide an opportunity to explore the general principles of adaptive evolution but are also effective approaches for identifying functional genes related to environmental adaptation, discovering domestication genes, and improving agronomic traits. Therefore, the comprehensive study of genomic convergence in plants should be a new frontier in evolutionary biology.

REFERENCES

Adamec, L. (1997). Mineral nutrition of carnivorous plants: a review. Bot. Rev. **63**:273–299.

Akashi, H., and Gojobori, T. (2002). Metabolic efficiency and amino acid composition in the proteomes of Escherichia coli and Bacillus subtilis. Proc. Natl. Acad. Sci. U S A **99**:3695–3700.

Argueso, J.L., Westmoreland, J., Mieczkowski, P.A., Gawel, M., Petes, T.D., and Resnick, M.A. (2008). Double-strand breaks associated with repetitive DNA can reshape the genome. Proc. Natl. Acad. Sci. U S A **105**:11845–11850.

Arnold, B.J., Lahner, B., DaCosta, J.M., Weisman, C.M., Hollister, J.D., Salt, D.E., Bomblies, K., and Yant, L. (2016). Borrowed alleles and convergence in serpentine adaptation. Proc. Natl. Acad. Sci. U S A **113**:8320–8325.

Besnard, G., Muasya, A.M., Russier, F., Roalson, E.H., Salamin, N., and Christin, P.A. (2009). Phylogenomics of C_4 photosynthesis in

sedges (Cyperaceae): multiple appearances and genetic convergence. Mol. Biol. Evol. **26**:1909–1919.

Bianconi, M.E., Dunning, L.T., Moreno-Villena, J.J., Osborne, C.P., and Christin, P.A. (2018). Gene duplication and dosage effects during the early emergence of C 4 photosynthesis in the grass genus Alloteropsis. J. Exp. Bot. **69**:1967–1980.

Brady, K.U., Kruckeberg, A.R., and Bradshaw, H.D. (2005). Evolutionary ecology of plant adaptation to serpentine soils. Annu. Rev. Ecol. Evol. Syst. **36**:243–266.

Castoe, T.A., de Koning, A.P.J., Kim, H.-M., Gu, W., Noonan, B.P., Naylor, G., Jiang, Z.J., Parkinson, C.L., and Pollock, D.D. (2009). Evidence for an ancient adaptive episode of convergent molecular evolution. Proc. Natl. Acad. Sci. U S A **106**:8986–8991.

Chao, D.Y., Dilkes, B., Luo, H., Douglas, A., Yakubova, E., Lahner, B., and Salt, D.E. (2013). Polyploids exhibit higher potassium uptake and salinity tolerance in Arabidopsis. Science **341**:658–659.

Chen, L., DeVries, A.L., and Cheng, C.H. (1997). Convergent evolution of antifreeze glycoproteins in Antarctic notothenioid fish and Arctic cod. Proc. Natl. Acad. Sci. U S A **94**:3817–3822.

Christin, P.A., Salamin, N., Savolainen, V., Duvall, M.R., and Besnard, G. (2007). C 4 photosynthesis evolved in grasses via parallel adaptive genetic changes. Curr. Biol. **17**:1241–1247.

Christin, P.A., Salamin, N., Muasya, A.M., Roalson, E.H., Russier, F., and Besnard, G. (2008). Evolutionary switch and genetic convergence on *rbcL* following the evolution of C 4 photosynthesis. Mol. Biol. Evol. **25**:2361–2368.

Christin, P.A., Weinreich, D.M., and Besnard, G. (2010). Causes and evolutionary significance of genetic convergence. Trends Genet. **26**:400–405.

Christin, P.A., Edwards, E.J., Besnard, G., Boxall, S.F., Gregory, R., Kellogg, E.A., Hartwell, J., and Osborne, C.P. (2012). Adaptive evolution of C 4 photosynthesis through recurrent lateral gene transfer. Curr. Biol. **22**:445–449.

Conant, G.C., and Wolfe, K.H. (2008). Turning a hobby into a job: how duplicated genes find new functions. Nat. Rev. Genet. **9**:938–950.

Conn, C.E., Bythell-douglas, R., Neumann, D., Yoshida, S., Whittington, B., Westwood, J.H., Shirasu, K., Bond, C.S., Dyer, K.A., and Nelson, D.C. (2015). Convergent evolution of strigolactone perception enabled host detection in parasitic plants. Science **349**:540–543.

Dassanayake, M., Haas, J., Bohnert, H., and Cheeseman, J. (2009). Shedding light on an extremophile lifestyle through transcriptomics. New Phytol. **183**:764–775.

Dunning, L.T., and Christin, P.A. (2020). Reticulate evolution, lateral gene transfer, and innovation in plants. Am. J. Bot. **107**:541–544.

Edwards, E.J., and Ogburn, R.M. (2012). Angiosperm responses to a low-CO 2 world: CAM and C 4 photosynthesis as parallel evolutionary trajectories. Int. J. Plant Sci. **173**:724–733.

Feng, X., Xu, S., Li, J., Yang, Y., Chen, Q., Lyu, H., and Zhong, C. (2020). Molecular adaptation to salinity fluctuation in tropical intertidal environments of a mangrove tree Sonneratia alba. BMC Plant Biol. **20**:1–14.

Foote, A.D., Liu, Y., Thomas, G.W.C., Vinar, T., Alfoldi, J., Deng, J., Dugan, S., van Elk, C.E., Hunter, M.E., Joshi, V., et al. (2015). Convergent evolution of the genomes of marine mammals. Nat. Genet. **47**:272–275.

Freeling, M. (2009). Bias in plant gene content following different sorts of duplication: tandem, whole-genome, segmental, or by transposition. Annu. Rev. Plant Biol. **60**:433–453.

Fukushima, K., Fang, X., Alvarez-Ponce, D., Cai, H., Carretero-Paulet, L., Chen, C., Chang, T.H., Farr, K.M., Fujita, T., Hiwatashi, Y., et al. (2017). Genome of the pitcher plant *Cephalotus* reveals genetic changes associated with carnivory. Nat. Ecol. Evol. **1**:1–9.

Giri, C., Ochieng, E., Tieszen, L.L., Zhu, Z., Singh, A., Loveland, T., Masek, J., and Duke, N. (2011). Status and distribution of mangrove forests of the world using earth observation satellite data. Glob. Ecol. Biogeogr. **20**:154–159.

Goldstein, R.A., Pollard, S.T., Shah, S.D., and Pollock, D.D. (2015). Nonadaptive amino acid convergence rates decrease over time. Mol. Biol. Evol. **32**:1373–1381.

Gong, Z., Xiong, L., Shi, H., Yang, S., Herrera-Estrella, L.R., Xu, G., Chao, D.Y., Li, J., Wang, P.Y., Qin, F., et al. (2020). Plant abiotic stress response and nutrient use efficiency. Sci. China Life Sci. **63**:635–674.

Ha, M., Li, W.H., and Chen, Z.J. (2007). External factors accelerate expression divergence between duplicate genes. Trends Genet. **23**:162–166.

Hanada, K., Zou, C., Lehti-Shiu, M.D., Shinozaki, K., and Shiu, S.H. (2008). Importance of lineage-specific expansion of plant tandem duplicates in the adaptive response to environmental stimuli. Plant Physiol. **148**:993–1003.

He, Z., Zhai, W., Wen, H., Tang, T., Wang, Y., Lu, X., Greenberg, A.J., Hudson, R.R., Wu, C.I., and Shi, S. (2011). Two evolutionary histories in the genome of rice: the roles of domestication genes. PLoS Genet. **7**:1–10.

He, Z., Li, X., Yang, M., Wang, X., Zhong, C., Duke, N.C., Wu, C.I., and Shi, S. (2019). Speciation with gene flow via cycles of isolation and migration: insights from multiple mangrove taxa. Natl. Sci. Rev. **6**:275–288.

He, Z., Xu, S., Zhang, Z., et al. (2020a). Convergent adaptation of the genomes of woody plants at the land-sea interface. Natl. Sci. Rev **7**:978–993.

He, Z., Xu, S., and Shi, S. (2020b). Adaptive convergence at the genomic level—prevalent, uncommon or very rare? Natl. Sci. Rev. **7**:947–951.

Heyduk, K., Moreno-Villena, J.J., Gilman, I.S., Christin, P.A., and Edwards, E.J. (2019). The genetics of convergent evolution: insights from plant photosynthesis. Nat. Rev. Genet. **20**:485–493.

Hoekstra, F.A., Golovina, E.A., and Buitink, J. (2001). Mechanism of plant desiccation tolerance. Trends Plant Sci. **6**:431–438.

Hu, Y., Wu, Q., Ma, S., Ma, T., Shan, L., Wang, X., Nie, Y., Ning, Z., Yan, L., Xiu, Y., et al. (2017). Comparative genomics reveals convergent evolution between the bamboo-eating giant and red pandas. Proc. Natl. Acad. Sci. U S A **114**:1081–1086.

Jiao, Y., Wickett, N.J., Ayyampalayam, S., Chanderbali, A.S., Landherr, L., Ralph, P.E., Tomsho, L.P., Hu, Y., Liang, H., Soltis, P.S., et al. (2011). Ancestral polyploidy in seed plants and angiosperms. Nature **473**:97–100.

Kelly, L.J., Plumb, W.J., Carey, D.W., Mason, M.E., Cooper, E.D., Crowther, W., Whittemore, A.T., Rossiter, S.J., Koch, J.L., and Buggs, R.J.A. (2020). Convergent molecular evolution among ash species resistant to the emerald ash borer. Nat. Ecol. Evol. **4**:1116–1128.

Kuo, J., and Hartog, C. Den (2006). Seagrass morphology, anatomy, and ultrastructure. In Seagrasses: Biology, Ecology and Conservation, A. Larkum, R.J. Orth, and C.M. Duarte, eds. (Dordrecht: Springer Netherlands), pp. 51–87.

Lee, K.M., and Coop, G. (2019). Population genomics perspectives on convergent adaptation. Philos. Trans. R. Soc. Lond. B Biol. Sci. **374**:20180236.

Lee, H., Golicz, A.A., Bayer, P.E., Severn-Ellis, A.A., Chan, C.K.K., Batley, J., Kendrick, G.A., and Edwards, D. (2018). Genomic

comparison of two independent seagrass lineages reveals habitat-driven convergent evolution. J. Exp. Bot. **69**:3689–3702.

Lenser, T., and Theißen, G. (2013). Molecular mechanisms involved in convergent crop domestication. Trends Plant Sci. **18**:704–714.

Li, Y., Liu, Z., Shi, P., and Zhang, J. (2010). The hearing gene Prestin unites echolocating bats and whales. Curr. Biol. **20**:55–56.

Li, Z., Defoort, J., Tasdighian, S., Maere, S., Van de Peer, Y., and De Smet, R. (2016). Gene duplicability of core genes is highly consistent across all angiosperms. Plant Cell **28**:326–344.

Liang, S., Zhou, R., Dong, S., and Shi, S. (2008). Adaptation to salinity in mangroves: Implication on the evolution of salt-tolerance. Chin. Sci. Bull. **53**:1708–1715.

Lin, Z., Li, X., Shannon, L.M., Yeh, C.T., Wang, M.L., Bai, G., Peng, Z., Li, J., Trick, H.N., Clemente, T.E., et al. (2012). Parallel domestication of the *Shattering1* genes in cereals. Nat. Genet. **44**:720–724.

Liu, Y., Cotton, J.A., Shen, B., Han, X., Rossiter, S.J., and Zhang, S. (2010). Convergent sequence evolution between echolocating bats and dolphins. Curr. Biol. **20**:53–54.

Losos, J.B. (2011). Convergence, adaptation, and constraint. Evolution **65**:1827–1840.

Lyu, H., He, Z., Wu, C.I., and Shi, S. (2018). Convergent adaptive evolution in marginal environments: unloading transposable elements as a common strategy among mangrove genomes. New Phytol. **217**:428–438.

Ma, T., Wang, J., Zhou, G., Yue, Z., Hu, Q., Chen, Y., Liu, B., Qiu, Q., Wang, Z., Zhang, J., et al. (2013). Genomic insights into salt adaptation in a desert poplar. Nat. Commun. **4**:2797.

Martin, A., and Orgogozo, V. (2013). The loci of repeated evolution: a catalog of genetic hotspots of phenotypic variation. Evolution **67**:1235–1250.

Maruyama, K., Urano, K., Yoshiwara, K., Morishita, Y., Sakurai, N., Suzuki, H., Kojima, M., Sakakibara, H., Shibata, D., Saito, K., et al. (2014). Integrated analysis of the effects of cold and dehydration on rice metabolites, phytohormones, and gene transcripts. Plant Physiol. **164**:1759–1771.

Oh, D., Dassanayake, M., Bohnert, H.J., and Cheeseman, J.M. (2012). Life at the extreme: lessons from the genome. Genome Biol. **13**:1–9.

Oh, D.H., Hong, H., Lee, S.Y., Yun, D.J., Bohnert, H.J., and Dassanayake, M. (2014). Genome structures and transcriptomes signify niche adaptation for the multiple-ion-tolerant extremophyte *Schrenkiella parvula*. Plant Physiol. **164**:2123–2138.

Ohno, S. (1970). Evolution by Gene Duplication (New York: Springer-Verlag).

Oliver, M.J., Tuba, Z., and Mishler, B.D. (2000). The evolution of vegetative desiccation tolerance in land plants. Plant Ecol. **151**:85–100.

Oliver, M.J., Velten, J., and Mishler, B.D. (2005). Desiccation tolerance in bryophytes: a reflection of the primitive strategy for plant survival in dehydrating habitats? Integr. Comp. Biol. **45**:788–799.

Parker, J., Tsagkogeorga, G., Cotton, J.A., Liu, Y., Provero, P., Stupka, E., and Rossiter, S.J. (2013). Genome-wide signatures of convergent evolution in echolocating mammals. Nature **502**:228–231.

Paul, S., Bag, S.K., Das, S., Harvill, E.T., and Dutta, C. (2008). Molecular signature of hypersaline adaptation: insights from genome and proteome composition of halophilic prokaryotes. Genome Biol. **9**:R70.

Pfeiffer, P., Goedecke, W., and Obe, G. (2000). Mechanisms of DNA double-strand break repair and their potential to induce chromosomal aberrations. Mutagenesis **15**:289–302.

Preite, V., Sailer, C., Syllwasschy, L., Bray, S., Ahmadi, H., Krämer, U., and Yant, L. (2019). Convergent evolution in *Arabidopsis halleri* and *Arabidopsis arenosa* on calamine metalliferous soils. Philos. Trans. R. Soc. Lond. B Biol. Sci. **374**:20180243.

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. Mol. Plant **11**:414–428.

Rokas, A., and Carroll, S.B. (2008). Frequent and widespread parallel evolution of protein sequences. Mol. Biol. Evol. **25**:1943–1953.

Ruiz, M., Quiñones, A., Martínez-Cuenca, M.R., Aleza, P., Morillon, R., Navarro, L., Primo-Millo, E., and Martínez-Alcántara, B. (2016). Tetraploidy enhances the ability to exclude chloride from leaves in carrizo citrange seedlings. J. Plant Physiol. **205**:1–10.

Sage, R.F., Christin, P.A., and Edwards, E.J. (2011). The C_4 plant lineages of planet Earth. J. Exp. Bot. **62**:3155–3169.

Schuermann, D., Molinier, J., Fritsch, O., and Hohn, B. (2005). The dual nature of homologous recombination in plants. Trends Genet. **21**:172–181.

Selmecki, A.M., Maruvka, Y.E., Richmond, P.A., Guillet, M., Shoresh, N., Sorenson, A.L., De, S., Kishony, R., Michor, F., Dowell, R., et al. (2015). Polyploidy can drive rapid adaptation in yeast. Nature **519**:349–351.

Shi, S., Huang, Y., Zeng, K., Tan, F., He, H., Huang, J., and Fu, Y. (2005). Molecular phylogenetic analysis of mangroves: independent evolutionary origins of vivipary and salt secretion. Mol. Phylogenet. Evol. **34**:159–166.

Šmarda, P., Bureš, P., Horová, L., Leitch, I.J., Mucina, L., Pacini, E., Tichý, L., Grulich, V., and Rotreklová, O. (2014). Ecological and evolutionary significance of genomic GC content diversity in monocots. Proc. Natl. Acad. Sci. U S A **111**:E4096–E4102.

Stern, D.L. (2013). The genetic causes of convergent evolution. Nat. Rev. Genet. **14**:751–764.

Storz, J.F. (2016). Causes of molecular convergence and parallelism in protein evolution. Nat. Rev. Genet. **17**:239–250.

Tenaillon, O., Tenaillon, O., Rodríguez-Verdugo, A., Gaut, R.L., McDonald, P., Bennett, A.F., Long, A.D., and Gaut, B.S. (2012). The molecular diversity of adaptive convergence. Science **335**:457–461.

Thomas, G.W.C., and Hahn, M.W. (2015). Determining the null model for detecting adaptive convergence from genomic data: a case study using echolocating mammals. Mol. Biol. Evol. **32**:1232–1236.

Thorogood, C.J., Bauer, U., and Hiscock, S.J. (2018). Convergent and divergent evolution in carnivorous pitcher plant traps. New Phytol. **217**:1035–1041.

Tomlinson, P.B. (1986). The Botany of Mangroves (Cambridge: Cambridge University Press).

Van Buren, R., Wai, C.M., Pardo, J., Giarola, V., Ambrosini, S., Song, X., and Bartels, D. (2018). Desiccation tolerance evolved through gene duplication and network rewiring in *Lindernia*. Plant Cell **30**:2943–2958.

Van Buren, R., Pardo, J., Wai, C.M., Evans, S., and Bartels, D. (2019). Massive tandem proliferation of ELIPs supports convergent evolution of desiccation tolerance across land plants. Plant Physiol. **179**:1040–1049.

Van De Peer, Y., Mizrachi, E., and Marchal, K. (2017). The evolutionary significance of polyploidy. Nat. Rev. Genet. **18**:411–424.

Vanneste, K., Baele, G., Maere, S., and Van De Peer, Y. (2014). Analysis of 41 plant genomes supports a wave of successful genome duplications in association with the Cretaceous – Paleogene boundary. Genome Res. **32**:1334–1347.

Wójcik, M., Gonnelli, C., Selvi, F., Dresler, S., Rostański, A., and Vangronsveld, J. (2017). Metallophytes of serpentine and calamine

soils – their unique ecophysiology and potential for phytoremediation. Adv. Bot. Res. **83**:1–42.

Woodhouse, M.R., and Hufford, M.B. (2019). Parallelism and convergence in post-domestication adaptation in cereal grasses. Philos. Trans. R. Soc. Lond. B Biol. Sci. **374**:20180245.

Wu, H.J., Zhang, Z., Wang, J.Y., Oh, D.H., Dassanayake, M., Liu, B., Huang, Q., Sun, H.X., Xia, R., Wu, Y., et al. (2012). Insights into salt tolerance from the genome of *Thellungiella salsuginea*. Proc. Natl. Acad. Sci. U S A **109**:12219–12224.

Wu, S., Han, B., and Jiao, Y. (2019). Genetic contribution of paleopolyploidy to adaptive evolution in angiosperms. Mol. Plant **13**:59–71.

Wu, D.-D., Yang, C.-P., Wang, M.-S., Dong, K.-Z., Yan, D.-W., Hao, Z.-Q., Fan, S.-Q., Chu, S.-Z., Shen, Q.-S., Jiang, L.-P., et al. (2020a). Convergent genomic signatures of high-altitude adaptation among domestic mammals. Natl. Sci. Rev. **7**:952–963.

Wu, C.-I., Wang, G.-D., and Xu, S. (2020b). Convergent adaptive evolution—how common, or how rare? Natl. Sci. Rev. **7**:945–946.

Xu, S., He, Z., Guo, Z., Zhang, Z., Wyckoff, G.J., Greenberg, A., Wu, C.I., Shi, S., and Zhang, J. (2017a). Genome-wide convergence during evolution of mangroves from woody plants. Mol. Biol. Evol. **34**:1008–1015.

Xu, S., He, Z., Zhang, Z., Guo, Z., Guo, W., Lyu, H., Li, J., Yang, M., Du, Z., Huang, Y., et al. (2017b). The origin, diversification and adaptation of a major mangrove clade (Rhizophoreae) revealed by whole genome sequencing. Natl. Sci. Rev. **4**:721–734.

Yang, R., Jarvis, D.E., Chen, H., Beilstein, M.A., Grimwood, J., Jenkins, J., Shu, S.Q., Prochnik, S., Xin, M., Ma, C., et al. (2013). The reference genome of the halophytic plant Eutrema salsugineum. Front. Plant Sci. **4**:46.

Yang, P.M., Huang, Q.C., Qin, G.Y., Zhao, S.P., and Zhou, J.G. (2014). Different drought-stress responses in photosynthesis and reactive oxygen metabolism between autotetraploid and diploid rice. Photosynthetica **52**:193–202.

Yang, X., Hu, R., Yin, H., Jenkins, J., Shu, S., Tang, H., Liu, D., Weighill, D.A., Cheol Yim, W., Ha, J., et al. (2017). The *Kalanchoë* genome provides insights into convergent evolution and building blocks of crassulacean acid metabolism. Nat. Commun. **8**:1899.

Yeaman, S., Hodgins, K.A., Lotterhos, K.E., Suren, H., Nadeau, S., Degner, J.C., Nurkowski, K.A., Smets, P., Wang, T., Gray, L.K., et al. (2016). Convergent local adaptation to climate in distantly related conifers. Science **353**:1431–1433.

Zhang, J. (2003). Evolution by gene duplication: an update. Trends Ecol. Evol. **18**:292–298.

Zhang, J. (2006). Parallel adaptive origins of digestive RNases in Asian and African leaf monkeys. Nat. Genet. **38**:819–823.

Zhang, J., and Kumar, S. (1997). Detection of convergent and parallel evolution at the amino acid sequence level. Mol. Biol. Evol. **14**:527–536.

Zhang, J., and Nei, M. (1997). Accuracies of ancestral amino acid sequences inferred by the parsimony, likelihood, and distance methods. J. Mol. Evol. **44**:139–146.

Zhang, Z., Qu, C., Yao, R., Nie, Y., Xu, C., Miao, J., and Zhong, B. (2019). The parallel molecular adaptations to the Antarctic cold environment in two psychrophilic green algae. Genome Biol. Evol. **11**:1897–1908.

Zhang, Y.-H., Ravi, V., Qin, G., Dai, H., Zhang, H.-X., Han, F.-M., Wang, X., Liu, Y.-H., Yin, J.-P., Huang, L.-M., et al. (2020). Comparative genomics reveal shared genomic changes in syngnathid fishes and signatures of genetic convergence with placental mammals. Natl. Sci. Rev. **7**:964–977.

Zhen, Y., Aardema, M.L., Medina, E.M., Schumer, M., and Andolfatto, P. (2012). Parallel molecular evolution in an herbivore community. Science **337**:1634–1637.

Zhou, X.X., Cai, L.L., Fu, M.P., Hong, L.W., Shen, Y.J., and Li, Q.Q. (2016). Progress in the studies of vivipary in mangrove plants. Chin. J. Plant Ecol. **40**:1328–1343.

Zou, Z., and Zhang, J. (2015a). No genome-wide protein sequence convergence for echolocation. Mol. Biol. Evol. **32**:1237–1241.

Zou, Z., and Zhang, J. (2015b). Are convergent and parallel amino acid substitutions in protein evolution more prevalent than neutral expectations? Mol. Biol. Evol. **32**:2085–2096.