Fine-scale empirical data on niche divergence and homeolog expression patterns in an allopolyploid and its diploid progenitor species

Reiko Akiyama¹, Jianqiang Sun², Masaomi Hatakeyama¹,³,⁴, Heidi E. L. Lischer¹,⁴,⁵, Roman V. Briskine¹,³, Angela Hay⁶, Xiangchao Gan⁶, Miltos Tsiantis⁶, Hiroshi Kudoh⁷, Masahiro M. Kanaoka⁸, Jun Sese⁹,¹⁰,¹¹, Kentaro K. Shimizu¹,¹² and Rie Shimizu-Inatsugi¹

¹Department of Evolutionary Biology and Environmental Studies, University of Zurich; 2Winterthurstrasse 190, Zurich CH-8057, Switzerland; 3Research Center for Agricultural Information Technology, National Agriculture and Food Research Organization, 3-1-1 Kannondai, Tsukuba, Ibaraki 305-8517, Japan; 4Functional Genomics Center Zurich, Winterthurstrasse 190, Zurich CH-8057, Switzerland; 5Swiss Institute of Bioinformatics, Quartier Sorge – Batiment Genopode, Lausanne CH-1015, Switzerland; 6Interfaculty Bioinformatics Unit, University of Bern, Baltzerstrasse 6, Bern CH-3012, Switzerland; 7Department of Comparative Development and Genetics, Max Planck Institute for Plant Breeding Research, Carl-von-Linné-Weg 10, Köln 50829, Germany; 8Center for Ecological Research, Kyoto University, Hizano 2-509-3, Otsu 520-2113, Japan; 9Division of Biological Science, Graduate School of Science, Nagoya University, Furo-cho, Chikusa-ku, Nagoya 464-8602, Japan; 10Humanome Lab, Inc., L-HUB 3F-1-4, Shumonmaizui-cho, Shinjuku, Tokyo 162-0822, Japan; 11Artificial Intelligence Research Center, AIST, 2-3-26 Aomi, Koto-ku, Tokyo 135-0064, Japan; 12AIST-Tokyo Tech RWBC-OIL, 2-12-1 Okayama, Meguro-ku, Tokyo 152-8550, Japan; 13Kihara Institute for Biological Research (KIBIR), Yokohama City University, 641-12 Maioka,Totsuka-ward, Yokohama 244-0813, Japan

Author for correspondence:
Rie Shimizu-Inatsugi
Email: rie.inatsugi@ieu.uzh.ch
Kentaro K. Shimizu
Email: kentaro.shimizu@ieu.uzh.ch

Received: 17 February 2020
Accepted: 9 November 2020

New Phytologist (2021) 229: 3587–3601
doi: 10.1111/nph.17101

Key words: allopolyploid, Cardamine, homeolog expression, temporal fluctuation, transcriptome, water availability.

Summary

- Polyploidization is pervasive in plants, but little is known about the niche divergence of wild allopolyploids (species that harbor polyploid genomes originating from different diploid species) relative to their diploid progenitor species and the gene expression patterns that may underlie such ecological divergence. We conducted a fine-scale empirical study on habitat and gene expression of an allopolyploid and its diploid progenitors.
- We quantified soil properties and light availability of habitats of an allotetraploid Cardamine flexuosa and its diploid progenitors Cardamine amara and Cardamine hirsuta in two seasons. We analyzed expression patterns of genes and homeologs (homeologous gene copies in allopolyploids) using RNA sequencing.
- We detected niche divergence between the allopolyploid and its diploid progenitors along water availability gradient at a fine scale: the diploids in opposite extremes and the allopolyploid in a broader range between diploids, with limited overlap with diploids at both ends. Most of the genes whose homeolog expression ratio changed among habitats in C. flexuosa varied spatially and temporally.
- These findings provide empirical evidence for niche divergence between an allopolyploid and its diploid progenitor species at a fine scale and suggest that divergent expression patterns of homeologs in an allopolyploid may underlie its persistence in diverse habitats.

Introduction

Polyploidization is pervasive in plants (Otto & Whitton, 2000) and is one of the major mechanisms of speciation (Ramsey & Schmske, 1998; Soltis & Soltis, 2009). In allopolyploids (species with multiple genome sets from different diploid species), differential expression of homeologs (potentially redundant copies of genes derived from the progenitor species) has been speculated to reflect selective use of the homeologs in responding to various environmental conditions (Madlung, 2013; Yoo et al., 2013), in addition to higher evolvability due to redundancy (Douglas et al., 2015; Paape et al., 2018). This, in turn, might enable allopolyploids to thrive in a range of environments. Consistent with this idea, manipulative experiments in laboratories demonstrated homeolog expression ratio changes in response to environmental factors such as temperature, water availability, and metal concentration (Liu & Adams, 2007; Dong & Adams, 2011; Akama et al., 2014; Liu et al., 2015; Paape et al., 2016; but see Combes et al., 2013). Recent empirical studies show that a small proportion (c. 1%) of all genes exhibits a ratio change, many of which are known to be involved in response to the specific conditions that were manipulated in the experiment (Akama et al., 2014; Paape et al., 2016). These results suggest that allopolyploids selectively express the homeologs relevant for an appropriate response to the environment. Although the expression patterns in the laboratory and in naturally fluctuating environments, or in...
natura, may be distinct (Shimizu et al., 2011; Yamasaki et al., 2017; Song et al., 2018), little is known about environment-dependent changes in the homeolog expression ratio of allopolyploids in natural habitats and about the divergence of expression patterns from progenitors.

Ecological differentiation enables allopolyploids to establish and persist amid well-established and possibly locally adapted diploid progenitor species, otherwise they would be excluded by competition (Hardin, 1960). Newly emerged polyploids may be prepared to exploit novel habitats, as they inherited divergent genomes of the diploid progenitor species (Levin, 2003; Paape et al., 2020). This leads to an expectation that allopolyploids and their progenitor species occupy different habitats (niche divergence). Niche divergence between an allopolyploid and its progenitor species has been extensively studied using ecological niche modeling, where species geographic distributions are modeled in relation to large-scale climatic and topographic variables (Wake et al., 2009). There is support for niche divergence, with intermediate and broad niche detected for allopolyploids, however, at the same time, several studies also report niche overlap between allopolyploids and their progenitors (Soltis et al., 2010; McIntyre, 2012; te Beest et al., 2012; Glennon et al., 2014; D. E. Soltis et al., 2014; P. S. Soltis et al., 2014; Marchant et al., 2016; Huynh et al., 2020; López-Alvarez et al., 2020). A possible explanation for niche overlap is the spatial scale of a study. Ecological niche modeling approaches typically evaluate environmental factors at a resolution of >1 km². Multiple environmental factors are known to vary at fine spatial scale, including water availability, light availability, and soil properties (Denney et al., 2020). These factors have been associated with niche divergence (Silvertown, 2004) and not yet incorporated in large-scale climatic modeling (www.worldclim.org). Under this circumstance, fine-scale empirical studies would have a better ability to detect niche divergence at fine scale. Furthermore, water availability varies not only spatially but also temporally (Denney et al., 2020). Data from multiple time points should enable the evaluation of temporal variation in habitats as well as differential gene expression (Aubin-Horth & Renn, 2009; Aikawa et al., 2010; Richards et al., 2012; Alvarez et al., 2015).

The genus Cardamine of Brassicaceae has long been a model to study ecological polyploid speciation (Howard, 1948; Hussein, 1948; Shimizu-Inatsugi et al., 2017; Akiyama et al., 2020; Sun et al., 2020). The allotetraploid Cardamine flexuosa in Europe derived from diploid progenitor species Cardamine amara (genome size 217–273 Mb) and Cardamine hirsuta (genome size 198–225 Mb) (Johnston et al., 2005; Mandáková et al., 2013; Mandáková et al., 2014; Gan et al., 2016; in-house measurement) offers a promising system with which to study the niche divergence and associated homeolog expression of allopolyploids in comparison with the diploid progenitor species. Cardamine flexuosa is estimated to have emerged between 10 000 ka and 1 Ma (Mandáková et al., 2014). The three species are common in central Europe (Carlson et al., 2009), and their distributions largely overlap at the scale of 5 km² in Switzerland (Supporting Information Fig. S1; Lauber et al., 2012). However, anecdotal reports suggest niche divergence among species, as C. hirsuta plants are typically observed on roadsides and in ditches, C. amara on river banks and in wet woodlands, and C. flexuosa along forest roads (Urbanska & Landolt, 1978; Koch et al., 2003; Libová et al., 2006; Grime et al., 2007; Lauber et al., 2012; Tedder et al., 2015). In a manipulative laboratory experiment where the three species underwent drought, submergence, or fluctuation of the two, C. hirsuta performed the best in drought and worst in submergence, the opposite was the case for C. amara, and C. flexuosa performed similarly well in all treatments (Shimizu-Inatsugi et al., 2017). The three species thus appear to show niche divergence along hydrological gradients, characterized by two distinct stresses – water-logging and drought – at the two ends of the gradient at a scale smaller than 5 km². Microarray analyses from the laboratory experiment showed that the gene expression induction pattern of C. flexuosa was similar to that of C. hirsuta under drought and to that of C. amara under submergence (Shimizu-Inatsugi et al., 2017). It is yet to be shown how the gene expression pattern of C. flexuosa compares to that of the diploid progenitor species in field. Additionally, a major limitation of the usage of the Arabidopsis microarray was that it allowed the quantification of only the sum of the expression levels of two homeologs of only c. 46% of the entire genes. Recently, HomEORoQ, a tool for analyzing the homeolog expression ratio change in allopolyploids (Akama et al., 2014; Kuo et al., 2020), and the genome sequence of C. hirsuta (Gan et al., 2016) became available, enabling evaluation of the expression of homeologs across genomes of C. flexuosa. These technical advances and findings from the laboratory experiment make C. flexuosa an allopolyploid with distinct ecologies and genomic tools. Such a study system is still rate (Soltis et al., 2016), and thus C. flexuosa offers a unique opportunity for studying ecological transcriptomics of an allopolyploid and its diploid progenitors. However, quantitative data from natural habitats are still lacking.

Here, we conducted a fine-scale empirical study over two seasons to quantify water availability, soil properties, and light availability in the habitats of C. flexuosa, C. amara and C. hirsuta within their native range in Switzerland, in three areas in and around Zurich, and to analyze homeolog expression patterns of C. flexuosa in comparison with diploid progenitor. Based on previous studies, we hypothesized that the three species occur in different habitats; that is, the species differ in realized niche (Hutchinson, 1957). We expect that C. flexuosa would inhabit a wide water-availability gradient, including sites with fluctuating water availability, and that it would differentially express homeologs similarly to either of the diploid progenitor species depending on environments. We addressed the following questions: (1) What is the relative contribution of different environmental factors to the occurrence of the allopolyploid C. flexuosa as well as the two diploid progenitor species? What are the key environmental factors? (2) Along the gradients of environmental factors contributing to the occurrence of the species, is C. flexuosa in a distinct and/or broader range than C. amara and C. hirsuta? (3) How does the gene expression of C. flexuosa relate to that of the coexisting diploid progenitor species? (4) Is the gene (homeolog) expression pattern of C. flexuosa associated with environmental
factor identified as the key for niche divergence based on ecological data?

Materials and Methods

Study species

Cardamine amara L. is a perennial widespread across most of central Europe (Marhold, 1998; Lihová et al., 2000; Grime et al., 2007). Populations of C. amara in lowland areas in Switzerland (< 1000 m asl) are reported to consist of diploids (2n = 2x = 16), and autotetraploid populations of subsp. austriaca have been observed at higher altitudes of the inner alpine valleys (Marhold, 1999; Lihová et al., 2000; Marhold et al., 2002). Our study sites were in the areas inhabited by only diploids. Cardamine hirsuta L. is a diploid (2n = 2x = 16) native to Europe (Marhold, 1995; Lihová et al., 2006; Grime et al., 2007). It is a winter annual and its seeds cannot tolerate submergence (Yatsu et al., 2003). Cardamine flexuosa With. is an annual or perennial tetraploid (2n = 4x = 32) native to Europe (Lihová et al., 2006), which originated from the cross of C. hirsuta and C. amara (Grime et al., 2007; Mandáková et al., 2014). Recently, plants that had been identified as C. flexuosa in eastern Asia turned out to belong to a distinct species, Cardamine occulta (Lihová et al., 2006; Marhold et al., 2016), making the ecology of C. flexuosa in Europe yet to be explored.

When not obvious, C. flexuosa and C. hirsuta were identified using morphological keys (Post et al., 2011). The incidence of backcrossing is seemingly low, if at all possible, as we found none out of a total of 115 plants in the study sites screened for their ploidy level using flow cytometry (CyFlow Space; Sysmex Partec, Goerlitz, Germany; using CyStain UV Precise P reagents, Sysmex Partec). Specimens of the study species collected from the study sites have been donated to Zürich Herbaria, the herbarium of the University of Zurich and ETH Zurich, where they are registered with barcodes Z-000167685 to Z-000167691 for C. hirsuta, Z-000167692 to Z-000167697 for C. flexuosa, and Z-000167698 to Z-000167704 for C. amara.

Study sites

The study was conducted in the seminatural and anthropogenic areas of Irchel (IR), Wehrenbach (WBH), and Kiessnacht-Tobel (KT) in and around Zurich, Switzerland (Fig. 1a). In 2013, we selected 18 sites (IR1–9, WBH1–3, KT1–5, 7) from the three areas (Fig. 1b–d). WBH1, WBH2 and WBH3 are separated by a hedge and a paved road. The composition of the species and the numbers of individuals of each species per site are summarized in Table S1. The study sites in 2014 were the same as in 2013, except for site KT6, which was newly added, and IR9, which was excluded because of its disappearance upon changed land use.

Measurement of habitat environment

The measurement was conducted on 17 sites where the species occurred in both 2013 and 2014 (excluding IR9 and KT6; see Study sites section). The data contained nine environmental factors: weight-based soil water content, soil carbon (C) to nitrogen (N) ratio (C/N ratio), soil pH, nitrate (NO₃⁻) concentration in fresh soil, ammonium (NH₄⁺) concentration in fresh soil, NO₃⁻ concentration in incubated soil, NH₄⁺ concentration in incubated soil, sky openness at the beginning of the season, and sky openness at the end of the season. For the detailed method, see Soil properties and Sky openness in Method S1.

Principal component analysis on habitat

Out of the nine environmental factors measured, NO₃⁻ concentration in fresh soil was excluded from analysis because most of the values were zeros and thus were not appropriate for principal component analysis (PCA). We conducted PCA using software R v.3.3.3 (R Core Team, 2017), with the VEGAN package (Oksanen et al., 2017). To examine the extent of variance of habitat environmental conditions of each species, we calculated the root-mean-square error for PC1 values. For each species for each year, we subtracted the mean PC1 value from the PC1 value for each site, squared the outcome, summed the squared values to be divided by the number of sites, and then took the root of it.

Logistic regression based on data on species occurrence and environmental factors

To examine the contributions of environmental conditions to the occurrence of each species in each year, we performed logistic regression analyses using PROC LOGISTIC in SAS v.9.3 (SAS Institute Inc., Cary, NC, USA). With the presence or absence in a given site as response variable, we ran 255 logistic regressions consisting of all possible combinations of the eight of the nine environmental factors as explanatory variables. Sky openness at the end of the season was excluded from this analysis because of high correlation with sky openness at the start of the season (Table S2). Out of the 255 models examined, the model with the lowest Akaike information criterion (AIC) value was considered as the best one. We assessed the significance of the explanatory variables in the best models using the log-likelihood ratio test (Littell et al., 1996).

Soil moisture monitoring

For regular nondestructive data collection, we monitored soil moisture weekly during growing seasons in 2013 and 2014 using an electrical sensor instead of sampling soil at every census. See Method S1 for details.

Reference-guided assembly of C. amara genome and RNA sequencing

We used the published genome sequence of C. hirsuta (Gan et al., 2016) and our C. amara genome sequence constructed by reference-guided de novo assembly based on C. hirsuta genome according to Lischer & Shimizu (2017). For the genome sequences, one specific line of C. amara from Switzerland (lab
ID: WEH ATS1) was clonally propagated and used for DNA extraction. We sequenced 368 million paired-end reads (three independent libraries with different insertion sizes of 150–500 bp) and 548 million mate-pair reads (four independent libraries with different insertion sizes from 3–15 kb). To annotate genes on C. amara genome, RNA samples from eight organs (anther, filament, sepal, petal, flower bud, young flower bud, and leaf) of another individual from Switzerland (lab ID: 3-
19A) were sequenced. All sequence information used for assembly and annotation has been deposited in the DNA Data Bank of Japan (DDBJ; http://www.ddbj.nig.ac.jp) as DRA Accession ID: DRA006316, BioProject ID: PRJDB4989, BioSample ID: SAMD00098907. Details of the de novo genome assembly and gene annotation by AUGUSTUS (Stanke, 2014) are in Method S1.

To identify homologous genes of Arabidopsis thaliana, we applied a reciprocal best-hit method. We aligned gene regions of the assembled genome of C. amara onto A. thaliana gene regions (TAIR10: https://www.arabidopsis.org) using BLASTN+ v.2.2.30 (Altschul et al., 1990) with $1 \times 10^{-6}$ as a cutoff of E-value. max_target_seq was set at 5, yielding five hits that exceeded the cutoff E-value. Among those five, the A. thaliana gene with the smallest E-value and largest bit score was defined as the homologous gene of the C. hirsuta gene. We identified 29 458 homeolog pairs (Gene Set 1) annotated in both the C. amara genome and the C. hirsuta genome to be used for RNA sequencing (RNAseq) analysis.

### Tissue sample collection and RNA-seq

We selected three sampling sites in IR1, KT2 and KT5, where C. flexuosa coexisted with C. hirsuta, grew on its own, and coexisted with C. amara, respectively. Soil moisture was distinct for these sites (Fig. S2), allowing the comparison of the expression pattern of C. flexuosa among the dry site cohabited by one diploid, the intermediate site, and the wet site cohabited by the other diploid. Table S3 summarizes the growth stage of three individuals of each species at each sampling site and date of sampling, that are treated as biological replicates. Leaf tissue of average-sized C. flexuosa was collected from the same three individuals on 18 April, 2 May and 16 May in 2013. The samples of the diploid progenitor species were analyzed for only 2 May, as C. hirsuta dried up by 16 May and C. amara was too small for sampling on 18 April and was lost to flooding after 2 May (Tables S3, S4). At each site, one individual beside a spot for soil moisture monitoring was marked at the beginning of the season. The samplings were conducted between 10:00 h and 14:00 h on sunny days to reduce any variation introduced by nonquantified stimuli and circadian rhythm (Alvarez et al., 2015). The leaf samples were stored in a tube with RNAlater (Qiagen), brought back to the laboratory and kept at 4°C overnight. Once RNAlater infiltrated through leaves, the sample tubes were stored at −80°C until RNA extraction by RNeasy Plant Mini Kit (Qiagen). Total RNA was used for the library synthesis using TruSeq Stranded mRNA Library Prep Kit (Illumina, San Diego, CA, USA) and sequenced using an Illumina HiSeq2500 at the Functional Genomics Center Zurich. On average, 7.8 million reads per sample were obtained from 32 samples (Table S4). RNA-seq reads with low quality were trimmed with TRIMPOMATIC (v.0.36) (Bolger et al., 2014) and then mapped to the C. amara (described earlier herein) and C. hirsuta (Gan et al., 2016) genomes with STAR (v.2.5.3a; Dobin et al., 2013) to classify the orthologs in each progenitor genome by PCA, treating one ortholog/homeolog at each site (IR1 or KT5) on each date (18 April, 2 May, and 16 May) vs 18 April (18 April vs 2 May, 2 May vs 16 May, and 16 May vs 18 April) on each sampling date was tested using the HOMEOROQ pipeline (Akama et al., 2014), with false discovery rate < 0.05 as the threshold for significance. Among Gene Set 1, we only analyzed genes whose SD among the sum of the biological replicates of the pair of concern was $< 0.3$ to remove false positives, according to the standard setting of HOMEOROQ. We also excluded extremely poorly expressed genes (FPKM $< 0.2$) from the analysis. This left 18 855 genes after filtering by SD and FPKM and subject to analysis to detect which genes showed significant homeolog expression ratio change between site-pairs or date-pairs and which genes did not. The analysis was performed using the SUSHI framework (Hatakeyama et al., 2016). Dataset S1 summarizes the results.

At the next step, a Gene Ontology (GO) enrichment analysis of the biological process was conducted using topGO v.2.32.0 with the elim algorithm at a significance level of $P = 0.05$ (Alexa & Rahnenfuhrer, 2016) for each site-pair on each sampling date to examine which gene categories are overrepresented. Gene Set 1 was targeted as parent population. Dataset S2 shows the results of
all comparisons. GO terms with < 10 or > 500 annotated genes were excluded from interpretation.

Results

Habitat of the allopolyploid C. flexuosa and its diploid progenitor species

To characterize the habitat of the three species in this study, the relative contributions of soil properties and light availability to habitat were evaluated by PCA, which suggested three differences among the species. First, sites of two diploid progenitor species, C. hirsuta (red border encompassing red and orange points) and C. amara (blue border encompassing black and blue points), were separated (Fig. 2). The separation along the PC1 axis indicates that C. amara sites exhibited high soil water content and soil C/N ratios (Fig. 2; Fig. S4a,b) as well as a low degree of sky openness (Fig. 2; Fig. S5) compared with C. hirsuta sites. The high C/N ratios in C. amara sites likely reflect low N concentrations in fresh and incubated soils (Fig. 2; Fig. S4c–f). Second, in the PCA plot, sites of the allopolyploid C. flexuosa (green border encompassing orange, green, and blue points) overlapped with both diploid progenitor species, reflecting the coexisting sites. Importantly, the sites cohabited by C. flexuosa and C. hirsuta (orange dots) tend to be in the middle of the PC1 axis compared with sites with C. hirsuta alone (red dots). A similar tendency, although weaker, was found with C. amara (more pronounced in properties such as soil water content – see later and Fig. S4a).

Third, compared with diploid progenitor species, the value range for PC1 was wider for the C. flexuosa sites. This was supported by larger root-mean-square error for PC1 values, which shows the average variance over samples: C. flexuosa: 0.68 and 0.54 in 2013 and 2014, respectively; C. hirsuta: 0.49; C. amara: 0.27; the same value for both years for diploid progenitor species due to consistency in the sites of occurrence.

Three sites in the area of WBH (indicated with asterisks in Fig. 2) exemplify the importance of the fine-scale environments. One was occupied by only C. amara (black point, WBH1) and the other two by only C. hirsuta (red points, WBH2 and 3). Given these sites were in close vicinity, this separation would not have been detected at a resolution of > 1 km² (Fig. 1a,c).

By examining each environmental factor separately, we found that the extent of overlap of the habitat of C. flexuosa and its diploid progenitor species varied among factors (Figs S4, S5). Cardamine flexuosa was in an intermediate value range, e.g. for soil water content (C. amara 30–60%, C. flexuosa 20–40%, C. hirsuta 20–35%, Fig. S4a) and sky openness (C. amara 5–35%, C. flexuosa 5–40%, C. hirsuta 10–70%, Fig. S5), all of which contributed to PC1. Therefore, soil water content, soil C/N ratio, and light availability seemed to correlate with habitat differentiation among species, and the habitat of C. flexuosa generally overlapped with that of the two diploid progenitor species in the intermediate value range.

Relative contributions of environmental factors to the occurrence of the study species

In all species, the best model with the lowest AIC to predict the occurrence of the species included water content, C/N ratio, and sky openness at season start. Importantly, the absolute value of the regression coefficient was the largest for water content, for the diploid progenitor species, and for the allopolyploid when sites in which C. flexuosa occurred in 2013 were concerned (Table 1). These results indicate that the three species share environmental factors that determine habitat differentiation and that soil water content is the most influential determinant.

The directions of the regression coefficients for the three factors differed among species (Table 1). The occurrence of C. amara was associated positively with water content and C/N ratio and negatively with sky openness at the season’s start. By contrast, the occurrence of C. hirsuta was associated with the three factors in the opposite direction. In C. flexuosa, the
directions of these regression coefficients consisted of a combination of those of its diploid progenitor species. For water content, negative and positive values were observed in 2013 and 2014, respectively, coinciding with the exclusion and inclusion of sites with extreme soil moisture (filled and open arrows in Fig. 3b; Table 1). Furthermore, the values of the regression coefficients for water content and C/N ratio of C. flexuosa were intermediate between those for C. amara and C. hirsuta. Hence, these quantitative analyses support the difference of the three species in the PCA plot (Fig. 2), in which the habitat of the two diploid progenitor species are distinct, and that of the allopolyploid C. flexuosa partly overlapped with those of its progenitors at the intermediate range.

Temporal fluctuation of the soil moisture

Because the key environmental factor for determining niche divergence turned out to be water content, which likely fluctuated naturally, we examined whether temporal patterns of habitat water availability varied among species. In 2013, the habitat’s level of soil moisture for the diploid progenitor C. amara was high, in contrast to the other diploid progenitor C. hirsuta (Fig. 3a,c). At any given time, the level of soil moisture of sites inhabited by only C. amara (black points in the figure) and by only C. hirsuta (red points in the figure) did not overlap with each other. The habitat’s level of soil moisture for the allopolyploid C. flexuosa overlapped partly with either of the diploid progenitor species in the intermediate range, but not at the high or low ends (Fig. 3b). The level of soil moisture in 2014 was generally similar to that in 2013, although at the beginning of the season the levels of soil moisture of the C. amara and C. flexuosa habitats were lower than they were in 2013 (Fig. 3a,b,e,f). Furthermore, the soil moisture range throughout the season was largest for C. flexuosa in both 2013 and 2014 (Fig. 3d,h). Thus, the habitats of C. flexuosa had a broad range of soil moisture overlapping with two diploid progenitor species at the intermediate range. The observed niche divergence along soil moisture provided a basis for studying homeolog expression patterns of C. flexuosa in comparison with its diploid progenitor species along a natural environmental gradient.

PCA of homeolog expression of C. flexuosa and its diploid progenitor species

To compare the whole genome gene expression pattern in the allopolyploid and diploid progenitor species, we analyzed RNA-seq data of C. flexuosa at three times at three sites with contrasting soil moisture levels, together with the cohabiting diploid progenitor species, C. hirsuta at IR1 and C. amara at KT5, on 2 May (Fig. S2; Tables S3, S4). The number of reads classified into each subgenome was almost equal in all C. flexuosa samples (Fig. S3), indicating no clear expression bias at subgenomic level. When we treated the sum expression of two homeologs as gene expression and compared the pattern with that of the diploid progenitor species by PCA, the whole-genome pattern showed a loose cluster for each species, but the pattern was not clearly correlated with site (Fig. S6). We then examined the expression pattern of the subgenomes separately. Within C. flexuosa individuals, the two subgenomes were clearly divided into two large groups by PC1 with 30% contribution rate, and the diploid genomes were located close to each child subgenome of the allopolyploid (Fig. 4a). PC2 roughly grouped each subgenome by date with 13% contribution rate, from early (downside) to late (upside) date. In the PC2 and PC3 plots, each study site formed a loose group regardless of subgenomes (Fig. 4b). Similarly, the resemblance to the expression pattern of the progenitors was detected when the expression level of each homeolog of C. flexuosa was compared with that of the diploid progenitor species by clustering analysis (Fig. S7).

Genes with homeolog expression ratio changes between two homeologs and GO analysis

To examine the number of genes associated with habitat in the allopolyploid, we further analyzed the RNA-seq data of C. flexuosa to extract homeolog pairs with significant ratio change among habitats, assuming that they reflect the difference in key environmental conditions. We examined c. 20 000 homeolog pairs (Table S4; Dataset S1), which were nearly twice as many as the number of genes examined in the previous study (Shimizu-Inatsugi et al., 2017). The number of genes with homeolog

| Species (ploidy level) | Year | AIC | Water content | C/N | Sky openness at season start |
|------------------------|------|-----|---------------|-----|----------------------------|
|                        |      |     | \( \beta \pm SE \) | \( \chi^2 \) | \( \beta \pm SE \) | \( \chi^2 \) |
| C. amara (2x)          | 2013, 2014 | –2.34 | 0.11 ± 0.11 | 8.14** | 0.06 ± 0.06 | 7.53** |
| C. hirsuta (2x)        | 2013, 2014 | –1.98 | –0.10 ± 0.10 | 7.35** | –0.04 ± 0.05 | 6.86** |
| C. flexuosa (4x)       | 2013 | 6.38 | –0.08 ± 0.08 | 5.25** | –0.005 ± 0.04 | 6.25*  |
|                        | 2014 | 4.46 | 0.03 ± 0.06  | 6.32*  | –0.02 ± 0.04 | 6.02*  |
|                        | 2013 and 2014 | 3.23 | –0.06 ± 0.07 | 5.64*  | –0.02 ± 0.04 | 6.02*  |
|                        | 2013 or 2014 | 7.72 | 0.003 ± 0.06 | 5.56*  | –0.003 ± 0.04 | 6.32*  |

The site of occurrence was consistent in the two years for C. amara and for C. hirsuta. As C. flexuosa occurred in only one of the two years in some sites, we ran four models with different response variables: occurrence in 2013, occurrence in 2014, occurrence in both 2013 and 2014, and occurrence in either 2013 or 2014. Water content and carbon-to-nitrogen ratio (C/N) were measured for soil. For each model, Akaike information criterion (AIC), partial regression coefficient \( \beta \pm SE \) for each variable, and chi-square value \( \chi^2 \) and significance (*) are given. *, \( P < 0.05 \); **, \( P < 0.01 \); ***, \( P < 0.001 \).
expression ratio change differed among site-pair by date combinations and constituted 0.9–3.3% of all genes (Fig. 5a; Table S5). There was little overlap between site-pairs on a given date or between dates of a site-pair (Fig. 5c,d). Similarly, date-pair by site combinations constituted 0.2–3.0% of all genes, with little overlap between time points at a given site or between sites of a given date-pair (Fig. 5b,e,f; Table S6). The smallest number of genes was detected to be different between KT2 and KT5 on any date, and the biggest difference was found between IR1 and KT2 (Fig. 5a; Table S5). IR1 showed the largest difference among dates, coinciding with the largest temporal change of the soil moisture level at this site (Figs 5b, S2; Table S6).

The GO analyses of the genes with homeolog expression ratio change identified 15 to 72 GO terms enriched in each site-pair by date combination (Dataset S2). Among these GO terms, three terms, GO:0009414 (response to water deprivation), GO:0006979 (response to oxidative stress), and GO:0000302 (response to reactive oxidative species), were detected in the highest proportion; that is, in six site-pair by date combinations (Fig. 6; Table S7). GO:0006833 (water transport) was found in five combinations (Table S7).

Given that the genes included in GO:0009414 could be most relevant to the fluctuating soil moisture, we examined the expression pattern of the 32 genes in GO:0009414 (Table S8). Different homeologous pairs changed the ratio depending on sites and times (Fig. S8). To illustrate different patterns of expression changes, we visualized the expression levels of the 32 genes together with those of cohabiting diploids in IR1 and KT5 (Fig. S9). In the gene CARHR155190, a homologue of Arabidopsis RDUF1 responsible for the response to a plant hormone ABA (Kim et al., 2012), Fh (C. hirsuta homeolog of C. flexuosa) had low expression at KT5. In the genes CARHR073790 and CARHR261630, the expression levels of Fh and Fa were similar to those of the cohabiting diploid. Which of Fa and Fh was more expressed than the other differed between IR1 and KT5 in the genes CARHR261630 and CARHR000120, whereas this was not the case in most other genes. These results suggest no single dominant pattern of homeolog regulation in the natural environments.

Discussion

Fine-scale niche acquisition by the allopolyploid C. flexuosa by combining the legacy of the progenitors

We documented the various environmental factors of the habitats of the allopolyploid and its diploid progenitor species (Figs 2, 3,
S4, S5; Table 1). Together with previous growth experiments with water stress treatments (Shimizu-Inatsugi et al., 2017), these data suggest the following information on niche divergence of the two diploid progenitor species in Cardamine and the allopolyploid derived from them. First, the habitats of the two diploid species are distinct. The environmental conditions of the two species were separated in the PCA (Fig. 2), among which soil water content contributed the most to habitat differentiation (Table 1). The temporal data of water availability (Fig. 3) indicated that the habitats of C. amara and C. hirsuta were wet and dry, respectively. Second, the habitat of the allopolyploid was broad, as indicated by the larger root-mean-square error along PC1 of Fig. 2, but the overlap with the progenitors was limited to the intermediate range, and thus excluded extremely wet or dry conditions (Fig. 2; Table 1). The broad range itself does not guarantee that the allopolyploid establishes its own niche. In the case of C. flexuosa, the temporal monitoring of habitat water availability revealed high fluctuation (Fig. 3), suggesting that the adaptation to a broad range of conditions can be advantageous in acquiring robustness in fluctuating environments. Consistent with this observation, C. flexuosa grew similarly well under the submergence, drought, and the fluctuation of the two treatments in the laboratory experiment (Shimizu-Inatsugi et al., 2017). Given these data, the allopolyploid C. flexuosa seems to have obtained a new distinct niche with fluctuating water availability (Fig. 3), which is in accordance with a previous record on occurrence along water gradient in Britain (Grime et al., 2007).

Empirical information lies central in the current study. As such, the scale of our data set corresponds to what manual data collection in the field can cover and does not suffice for niche modeling. Despite the difference in approaches, there is an intriguing similarity in the findings from current and previous studies. In this study, PCA and multiple regression on empirical data of habitat at fine scale detected the allopolyploid occupying intermediate and broader environmental conditions. Previous studies using ecological niche modeling at continental scale reported intermediacy (Theodoridis et al., 2013; Harbert et al., 2014; Lopez-Alvarez et al., 2015; Marchant et al., 2016) or broader niche (McIntyre, 2012; Lopez-Alvarez et al., 2015; Marchant et al., 2016; Huynh et al., 2020). In contrast to previous studies, we observed species occurrence and habitat at small geographical scale. In some cases, the species segregated < 1 km² in a contrasting habitat; for example, C. amara in WBH1 vs C. hirsuta in WBH2 and 3 (Fig. 1). We did not capture this habitat differentiation at a larger geographic scale (IR, WBH, and KT; Fig. 1a). It is noteworthy that the significance of fine spatial scale has been recently highlighted in ecological niche modeling (Mertes & Jetz, 2018; Graham et al., 2019). Ecological niche modeling may also benefit from incorporating more diverse environmental factors. In this study, soil properties, light availability, and fluctuation of water availability during growing seasons were associated with the occurrence of the study species (Figs 2, 3, S4, S5 Table 1), confirming the importance of these factors in niche divergence (Silvertown, 2004). Overall, our results underscore that empirical data on environmental factors contributing to species occurrence at a fine spatial scale are informative.

Expression pattern of homeologs in the allopolyploid C. flexuosa associated with the key factor in niche divergence

The comparison between transcriptomes of the diploids and the allopolyploid has shown that the most dominant factors shaping expression patterns were genome type and date, and the effect by site was less dominant (Figs 4, S6). This suggests that most homeologs in the allotetraploid are strongly regulated by the legacy of progenitor genomes and subsequently by growing stage.
Fig. 5 The numbers and proportional Venn diagrams for the number of genes with homeolog expression ratio change in Cardamine flexuosa. (a, b) A summary of the number of genes with homeolog expression ratio change (a) between sites (IR, Irchel; KT, Küsnacht-Tobel) and (b) between dates among 23,182 homeologous pairs analyzed. Different sites and dates are indicated with different colors: IR1 in orange, KT2 in green, and KT5 in dark blue; light brown, 18 April; brown, 2 May; dark brown, 16 May. (c) The number of genes overlapping between site-pairs on each date. (d) The number of genes overlapping between dates for each site-pair. (e) The number of genes overlapping between date-pairs on each site. (f) The number of genes overlapping between sites for each date-pair. In (c–f), the size of circles with less than 190 genes is disproportionally enlarged to make numbers visible.
The allopolyploid species obtained a generalist niche by transcriptomic plasticity

A fundamental question on niche divergence is how the species coexist (Silvertown, 2004). We propose two mutually nonexclusive reasons for the coexistence of the trio of Cardamine species. First, unlike diploid progenitors, C. flexuosa can thrive in fluctuating environments, probably thanks to the transcriptomic plasticity achieved by the combination of two progenitor genomes specialized in contrasting environments (Madlung, 2013; Paape et al., 2016). The majority of the homeolog pairs did not change the ratio between site-pairs or time points in the current study; however, those with the ratio change showed little overlap across sites or over time (Fig. 5c–f; Tables S5, S6). This suggests that the allopolyploid flexibly expresses diverse homeologs at a given place or time in responding to fluctuating environment. Second, although C. flexuosa can occur in both wet and dry habitats, it may not be able to withstand extreme environments as much as the diploid progenitor species can. This can be attributed to the diploid progenitor species being better adapted to their respective niches compared with the allopolyploid, and thus the allopolyploid could not compete/displace its progenitor species from those niches (e.g. Lopez-Alvarez et al., 2015). It can also be due to attenuated response of the allopolyploid to environments at the transcriptomic level. Indeed, a previous transcriptomic study showed that the degree of upregulation of stress response genes was, in general, lower in C. flexuosa than in the advantageous diploid (Shimizu-Inatsugi et al., 2017). Thus, it seems that the allopolyploid C. flexuosa is a generalist, both ecologically and transcriptomically, filling the open niche between diploid progenitor species (Parisod & Broennimann, 2016). At the same time, becoming a generalist has a trade-off, in that its stress tolerance might also be attenuated, which may lead to C. flexuosa not being able to compete against progenitors in the extremes of the environmental gradient.
The generalist niche and transcriptomic plasticity of an allopolyploid we proposed for *C. flexuosa* can be common to other allopolyploid species. For example, the genus *Cardamine* has been proposed to have experienced adaptive radiation along water availability gradients by recurrent allopolyploidization, and the anecdotally described habitats of allopolyploid species including *Cardamine scutata* and *Cardamine insueta* are intermediate and/or fluctuating (Shimizu-Inatsugi *et al.*, 2017), consistent with a generalist niche. At the transcriptional level, genes involved in abiotic stress responses might be under cis-regulation in *C. flexuosa* and in other allopolyploid species, such as cotton, coffee, and *Arabidopsis* (Chaudhary *et al.*, 2009; Dong & Adams, 2011; Combes *et al.*, 2013; Yoo *et al.*, 2013; Carlson *et al.*, 2017). Both in the present study (Table S5) and previous studies on Brassicaceae in the laboratory (Akama *et al.*, 2014; Paape *et al.*, 2016), only a small portion (< 4%) of all genes exhibited the homeolog expression ratio change. Future studies may address the role of a limited number of distinctively regulated environment-response homeologs in the establishment and persistence of an allopolyploid in a generalist niche.

**Acknowledgements**

We thank G. Halstead-Nussloch, M. Helling, S. Milosavljevic, A. Morishima, T. Paape, S. Röthlisberger, K. K. S. Ng, A. Tedder, M. Yamazaki, N. Zweifel, A. M. Suarez, and other members of the Shimizu Group, D. Schneebeli, S. Moor, P. Enz, B. Hirzel, D. Schlagenhauf, N. Kofmehl, H. Bär, M. Freund, P. Niklaus, R. Husi, A. Ferrari, T. Zipfe, M. Furler, B. Schmid, P.C. Joerg, R. Leiterer, D. Markulin, B. Kellenberger, J. Sugisaka, L. Koyama, K. Fukushima, and N. Sletvold for their advice and assistance. We also thank J. Kühn-Georgijevic at the Functional Genomics Center of Zurich for RNA-seq analysis, the Botanical Garden of the University of Zurich and Gemeinde Küsnacht-Tobel for permits for field studies, Swisstopo for geographical data, and six anonymous reviewers for their comments. This study was funded by the Swiss National Science Foundation grant nos. 31003A_182318 and 31003A_159767 to RS-I and KKS; by University Research Priority Programs, Evolution in Action of the University of Zurich to RS-I and KKS; by Swiss National Science Foundation Marie-Heim Vögtlin grant no. PMPD3_134200 to RSI; by the Human Frontier Science Program to KKS, AH, and J Sese; by the Japan Science and Technology Agency, Core Research for Evolutionary Science and Technology grant no. JPMJCR1603 to KKS; by Indo-Swiss Collaboration in Biotechnology to MH; and by KAKENHI grant nos. 16H06469 to KKS and J Sese, 18H04785 to KKS, and 26113709 and 16H01463 to MMK and KKS.

**Author contributions**

RA, KKS and RS-I conceived the study, RA, HK and RS-I designed the study and collected the data, HELL, RVB, AH, XG, and MT provided the sequence data, RA, JSun, MH and JSese analyzed the data, RA, JSun, MH, MMK, KKS and RS-I wrote the manuscript with input from the other authors.

**ORCID**

Reiko Akiyama [https://orcid.org/0000-0002-6702-5101](https://orcid.org/0000-0002-6702-5101)  
Roman V. Briskine [https://orcid.org/0000-0002-6831-3914](https://orcid.org/0000-0002-6831-3914)  
Xiangchao Gan [https://orcid.org/0000-0001-6398-5191](https://orcid.org/0000-0001-6398-5191)  
Masami Hatakeyama [https://orcid.org/0000-0002-7409-9694](https://orcid.org/0000-0002-7409-9694)  
Hirosi Kudoh [https://orcid.org/0000-0001-9777-4886](https://orcid.org/0000-0001-9777-4886)  
Heidi E. L. Lischer [https://orcid.org/0000-0002-9616-2092](https://orcid.org/0000-0002-9616-2092)  
Jun Sese [https://orcid.org/0000-0003-3495-4382](https://orcid.org/0000-0003-3495-4382)  
Kentaro K. Shimizu [https://orcid.org/0000-0002-6483-1781](https://orcid.org/0000-0002-6483-1781)  
Rie Shimizu-Inatsugi [https://orcid.org/0000-0002-9899-591X](https://orcid.org/0000-0002-9899-591X)  
Jianqiang Sun [https://orcid.org/0000-0002-3438-3199](https://orcid.org/0000-0002-3438-3199)

**Data availability**

The sequence information used in this study will be openly available as of 31 December 2020 in the DDBJ (http://www.ddbj.nig.ac.jp) as DRA Accession ID: DRA006316, BioProject ID: PRJDB4989, BioSample ID: SAMD00098907 as well as DRA006314, BioProject ID: PRJDB4989, BioSample ID: SAMD00097398-SAMD00097406. The data that support the findings of this study are available in Dryad at https://doi.org/10.5061/dryad.4qrj6q8h

**References**

Aikawa S, Kobayashi MJ, Satake A, Shimizu KK, Kudoh H. 2010. Robust control of the seasonal expression of the *Arabidopsis* FLC gene in a fluctuating environment. *Proceedings of the National Academy of Sciences, USA* 107: 11632–11637.

Akama S, Shimizu-Inatsugi R, Shimizu KK, Sese J. 2014. Genome-wide quantification of homeolog expression ratio revealed nonstochastic gene regulation in synthetic allopolyploid *Arabidopsis*. *Nucleic Acids Research* 42: e46.

Akiyama R, Milosavljevic S, Leutenegger M, Shimizu-Inatsugi R. 2020. Trait-dependent resemblance of the flowering phenology and floral morphology of the allopolyploid *Cardamine flexuosa* to those of the parental diploids in natural habitats. *Journal of Plant Research* 133: 147–155.

Alexa A, Rahnenfuhrer J. 2016. *topGO*: enrichment analysis for gene ontology. R v.2.26.0. [WWW document] URL: https://bioconductor.org/packages/topGO/.

Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ. 1990. Basic local alignment search tool. *Journal of Molecular Biology* 215: 403–410.

Alvarez M, Schrey AW, Richards CL. 2015. Ten years of transcriptomics in wild populations: what have we learned about their ecology and evolution? *Molecular Ecology* 24: 710–725.

Ashin-Horth N, Renn SC. 2009. Genomic reaction norms: using integrative biology to understand molecular mechanisms of phenotypic plasticity. *Molecular Ecology* 18: 3763–3780.

Bolger AM, Lohse M, Usadel B. 2014. TRIMMOMATIC – a flexible trimmer for Illumina sequence data. *Bioinformatics* 30: 2114–2120.

Carlsen T, Bleeker W, Hurka H, Elven R, Brochmann C. 2009. Biogeography and phylogeny of *Cardamine* (Brassicaceae). *Annals of the Missouri Botanical Garden* 96: 215–236.
Carlson KD, Fernandez-Pozo N, Bombarely A, Pisupati R, Mueller LA. Madlung A. 2017. Natural variation in stress response gene activity in the allopolyploid Arabidopsis sieberi. BMC Genomics 18: e653.

Chaudhary B, Flagel L, Stupar RM, Udall JA, Verma N, Springer NM, Wendel JF. 2009. Reciprocal silencing, transcriptional bias and functional divergence of homoeology in polyploid cotton (Gossypium). Genetics 182: 503–517.

Combres MC, Deregue A, Severac D, Bertrand B, Lashermes P. 2013. Contribution of subgenomes to the transcriptome and their intertwined regulation in the allopolyploid Coffea arabica grown at contrasted temperatures. New Phytologist 200: 251–260.

Denney DA, Jameel MI, Bemmels JB, Rochford ME, Anderson JT. 2020. Small spaces, big impacts: contributions of micro-environmental variation to population persistence under climate change. Aob Plants 12: pla005.

Dobin A, Davis CA, Schlesinger D, Frenkow J, Zaleski C, Jha S, Batut P, Chaisson M, Gingeras TR. 2013. STAR: ultrastar universal RNA-seq aligner. Bioinformatics 29: 15–21.

Dong S, Adams KL. 2011. Differential contributions to the transcriptome of duplicated genes in response to abiotic stresses in natural and synthetic polyploids. New Phytologist 190: 1045–1057.

Douglas GM, Gos G, Steige KA, Salcedo A, Holm K, Josephs EB, Arunkumar R, Agren JA, Hazzouri KM, Wang W et al. 2015. Hybrid origins and the earliest stages of diploidization in the highly successful recent polyploid Capsella bursa-pastoris. Proceedings of the National Academy of Sciences, USA 112: 2806–2811.

Gan X, Hay A, Kwantes M, Haberer G, Hallab A, Ioio RD, Hofhuis H, Pieper B, Cartolano M, Neumann U et al. 2016. The Cardamine hirumata genome offers insight into the evolution of morphological diversity. Nature Plants 2: e16176.

Glennon KL, Ritchie ME, Segraves KA. 2014. Evidence for shared broad-scale climatic niches of diploid and polyploid plants. Ecology Letters 17: 574–582.

Graham LJ, Spake R, Gillings S, Watts K, Eigenbord F. 2019. Incorporating fine-scale environmental heterogeneity into broad-extent models. Methods in Ecology and Evolution 10: 767–778.

Grime JP, Hodgson JG, Hunt R. 2007. Comparative plant ecology: a functional approach to common British species. Colvend, UK: Castlepoint Press.

Harbert RS, Brown AHD, Doyle JJ. 2014. Climate niche modeling in the perennial Glycine (Leguminosae) allopolyploid complex. American Journal of Botany 101: 710–721.

Hardin G. 1960. The competitive exclusion principle. Science 131: 1292–1297.

Hatakeyama M, Opitz L, Russo G, Qi W, Schlapbach R, Rehrauer H. 2016. SSEI: an exquisite recipe for fully documented, reproducible and reusable NGS data analysis. BMC Bioinformatics 17: e228.

Howard HW. 1948. Chromosome number of Cardamine pratensis. Nature 161: 277.

Hussein F. 1948. Chromosome number of Cardamine pratensis. Nature 161: 1015.

Hutchinson GE. 1957. Population studies – animal ecology and demography – concluding remarks. Cold Spring Harbor Symposium on Quantitative Biology 22: 415–427.

Huynh S, Broennimann O, Guisan A, Felber F, Parisod C. 2020. Contribution of subgenomes to the transcriptome and their intertwined regulation in the allopolyploid Coffea arabica grown at contrasted temperatures. New Phytologist 200: 251–260.

Hutchinson GE. 1957. Population studies – animal ecology and demography – concluding remarks. Cold Spring Harbor Symposium on Quantitative Biology 22: 415–427.

Huynh S, Brønnimann O, Guisan A, Felber F, Parissod C. 2020. Genetic evolutionary history of polyploid species. Ecology Letters 23: 663–673.

Johnston JS, Pepper AE, Hall AE, Chen ZJ, Hodnett G, Drabek J, Lopez R, Price HJ. 2005. Evolution of genome size in Brassicaceae. Annales of Botany 95: 229–235.

Kim SJ, Ryu MY, Kim WT. 2012. Suppression of Arabidopsis RING-UDF1117 E3 ubiquitin ligases, AtRUDUF1 and AtRUDUF2, reduces tolerance to ABA-mediated drought stress. Biochemical and Biophysical Research Communications 420: 141–147.

Koch M, Huthmann M, Bernhardt K-G. 2003. Cardamine amara L. (Brassicaceae) in dynamic habitats: genetic composition and diversity of seed bank and established populations. Basic and Applied Ecology 4: 339–348.

Kuo TC, Hatakeyama M, Tameshige T, Shimizu KK, Sese J. 2020. Homeolog expression quantification methods for allopolyploids. Briefings in Bioinformatics 21: 395–407.

Lauber K, Wagner G, Gygax A. 2012. Flora Helvetica. Berne, Switzerland: Haupt Verlag AG.

Levin DA. 2003. The ecological transition in speciation. New Phytologist 161: 91–96.

Li Y, Zhu Y, Liu Y, Shu Y, Meng F, Lu Y, Bai X, Liu B, Guo D. 2008. Genome-wide identification of osmotic stress response gene in Arabidopsis thaliana. Genomics 92: 488–493.

Liao Y, Smyth GK, Shi W. 2014. FeatureCounts: an efficient general purpose program for assigning sequence reads to genomic features. Bioinformatics 30: 923–930.

Lihová J, Marhold K, Kudoh H, Koch MA. 2006. Worldwide phylogeny and biogeography of Cardamine flexuosa (Brassicaceae) and its relatives. American Journal of Botany 93: 1206–1221.

Lihová J, Marhold K, Neuffer B. 2000. Taxonomy of Cardamine amara (Cruciferae) in the Iberian Peninsula. Taxon 49: 747–763.

Lischer HEL, Shimizu KK. 2017. Reference-guided de novo assembly approach improves genome reconstruction for related species. BMC Bioinformatics 18: e474.

Littell RC, Milliken GA, Stroup WW, Wolfinger RD. 1996. SAS system for mixed models. Cary, NC, USA: SAS Institute.

Liu Z, Adams KL. 2007. Expression partitioning between genes duplicated by polyploidy under abiotic stress and during organ development. Current Biology 17: 1669–1674.

Liu Z, Xin M, Qin J, Peng H, Ni Z, Yao Y, Sun Q. 2015. Temporal transcriptome profiling reveals expression partitioning of homeologous genes contributing to heat and drought acclimation in wheat (Triticum aestivum L.). BMC Plant Biology 15: e152.

López-Alvarez D, Manzaneda AJ, Rey PJ, Giraldo P, Benavente E, Allainiguillaume J, Murl L, Caicedo AL, Hazen SP, Breiman A et al. 2015. Environmental niche variation and evolutionary diversification of the Brachypodium distachyon grass complex species in their native circum-Mediterranean range. American Journal of Botany 102: 1073–1088.

Lowry DB, Logan TL, Santuari L, Hardtke CS, Richards JH, DeRose-Wilson LJ, McKay JK, Sen S, Juenger TE. 2013. Expression quantitative trait locus mapping across water availability environments reveals contrasting associations with genomic features in Arabidopsis. The Plant Cell 25: 3566–3579.

Madlung A. 2013. Polyploidy and its effect on evolutionary success: old questions revisited with new tools. Heredity 110: 99–104.

Mandáková T, Kovářík A, Zozomová-Líhová J, Shimizu-Intasugi R, Shimizu KK, Mummnehoff K, Marhold K, Lysak MA. 2013. The more the merrier: recent hybridization and polyploidy in Cardamine. The Plant Cell 25: 3280–3295.

Mandáková T, Marhold K, Lysak MA. 2014. The widespread crucifer species Cardamine flexuosa is an allotetraploid with a conserved subgenomic structure. New Phytologist 201: 982–992.

Marchant DB, Soltis DE, Soltis PS. 2016. Patterns of abiotic niche shifts in allopolyploids relative to their progenitors. New Phytologist 212: 708–718.

Marhold K. 1995. Taxonomy of the genus Cardamine L. (Cruciferae) in the Carpathians and in Pannonia. III. Folia Geobotanica et Phyto-taxonomica 30: 397–434.

Marhold K. 1998. Morphometric comparison of diploid populations of Cardamine amara (Brassicaceae) from central Europe and the Balkan Peninsula. Thaizia – Journal of Botany 8: 19–32.

Marhold K. 1999. Taxonomic evaluation of the tetraploid populations of Cardamine amara (Brassicaceae) from the eastern Alps and adjacent areas. Botanica Helvetica 109: 67–84.

Marhold K, Huthmann M, Hurka H. 2002. Evolutionary history of the polyploid complex of Cardamine amara (Brassicaceae): isozyme evidence. Plant Systematics and Evolution 233: 15–28.

Marhold K, Slender M, Kudoh H, Zozomova-Líhová J. 2016. Cardamine occulta; the correct species name for invasive Azian plants previously classified as C. flexuosa, and its occurrence in Europe. Phytokeys 62: 57–72.

McIntyre PJ. 2012. Polyploidy associated with altered and broader ecological niches in the Claytonia perfoliata (Portulacaceae) species complex. American Journal of Botany 99: 655–662.

Mertes K, Jetz W. 2018. Disentangling scale dependencies in species environmental niches and distributions.Ecography 41: 1604–1615.

Oksanen J, Guillaume Blanchet F, Friendly M, Kindt R, Legendre P, McGlinn D, Minchin PR, O’Hara RB, Simpson GL, Solymos P et al. 2017. Vegan:
community ecology package. R v.2.4.4. [WWW document] URL https://cran.r-project.org/web/packages/vegan/index.html [accessed 31 October 2017].

Otto SP, Whitton J. 2000. Polyploidy incidence and evolution. Annual Review of Genetics 34: 401–437.

Paepe T, Hatakeyama M, Shimizu-Inatsugi R, Cereghetti T, Onda Y, Kenata T, Sese J, Shimizu KK. 2016. Conserved but attenuated parental gene expression in allopolyploids: constitutive zinc hyperaccumulation in the allooctoploid Arabidopsis kochiachatra. Molecular Biology and Evolution 33: 2781–2800.

Paepe T, Briskine RV, Halstead-Nussloch G, Lischer HEL, Shimizu-Inatsugi R, Hatakeyama M, Kenta T, Nishiyama T, Sabirov R, Sese J et al. 2018. Patterns of polymorphism and selection in the subgenomes of the allopolyploid Arabidopsis kochiachatra. Nature Communications 9: e3909.

Paepe T, Akiyama R, Cereghetti T, Onda Y, Hirao A, Teta T, Shimizu KK. 2020. Experimental and field data support range expansion in an allopolyploid Arabidopsis owing to parental legacy of heavy metal hyperaccumulation. Frontiers in Genetics 11: e565854.

Pariisod C, Broennimann O. 2016. Towards unified hypotheses of the impact of polyploidy on ecological niches. New Phytologist 212: 540–542.

Post AR, Krings A, Xiang J, Sosinski BR, Neal JC. 2011. Molecular basis of flowering under natural long-day conditions in Cardamine hirsuta. Molecular Plant 4: 3587–3601.

Post AR, Krings A, Xiang J, Sosinski BR, Neal JC. 2011. Molecular basis of flowering under natural long-day conditions in Cardamine hirsuta. Molecular Plant 4: 3587–3601.

Pysek P. 2012. The role of hybridization in plant speciation. Annual Review of Plant Biology 63: 401–427.

Research. Tedder A, Helling M, Pannell JR, Shimizu-Inatsugi R, Kawagoe T, van Campen J, Sese J, Shimizu KK. 2015. Female fertility associated with increased clonal propagation suggests a unique combination of androdioecy and asexual reproduction in populations of Cardamine amara (Brassicaceae). Annals of Botany 115: 763–776.

Soltis PS, Soltis DE. 2009. The role of hybridization in plant speciation. Annual Review of Plant Biology 60: 561–588.

Soltis PS, Soltis DE. 2009. The role of hybridization in plant speciation. Annual Review of Plant Biology 60: 561–588.

Soltis DE, Buggs RJA, Doyle JJ, Soltis PS. 2010. What we still don’t know about polyploidy. Taxon 59: 1387–1403.

Soltis DE, Visger CJ, Soltis PS. 2014. The polyploidy revolution then . . . and now; Stebbins revisited. American Journal of Botany 101: 1057–1078.

Soltis PS, Liu X, Marchant DB, Visger CJ, Soltis DE. 2014. Polyploidy and novelty: Gottlieb’s legacy. Philosophical Transactions of the Royal Society B: Biological Sciences 369: e20130351.

Soltis DE, Visger CJ, Marchant DB, Soltis PS. 2016. Polyploidy: pitfalls and paths to a paradigm. American Journal of Botany 103: 1146–1166.

Song YH, Kubota A, Kwon MS, Covington MF, Lee N, Taagen ER, Cintrón DL, Hwang DY, Akiyama R, Hodge SK et al. 2018. Molecular basis of flowering under natural long-day conditions in Arabidopsis. Nature Plants 4: 824–835.

Staneke C. 2014. Incorporating RNA-seq into AUGUSTUS. [WWW document] URL https://bioinf.uni-greifswald.de/augustus/binaries/readme.rnaseq.html [accessed 15 January 2015].

Sun J, Shimizu-Inatsugi R, Hofhuis H, Shimizu K, Hay A, Shimizu KK. 2020. A recently formed triploid Cardamine insueta inherits leaf vivipary and submergence tolerance traits of parents. Frontiers in Genetics 11: e567262.

tc Beest M, Le Roux J, Richardson DM, Brysting AK, Suda J, Kubesová M, Pysek P. 2012. The more the better? The role of polyploidy in facilitating plant invasions. Annals of Botany 109: 19–45.

Tedder A, Helling M, Pannell JR, Shimizu-Inatsugi R, Kawagoe T, van Campen J, Sese J, Shimizu KK. 2015. Female fertility associated with increased clonal propagation suggests a unique combination of androdioecy and asexual reproduction in populations of Cardamine amara (Brassicaceae). Annals of Botany 115: 763–776.

Theodoridis S, Randin C, Broennimann O, Patsiou T, Conti E. 2013. Divergent and narrower climatic niches characterize polyploidy species of European primroses Primula sect. Aleuritia. Journal of Biogeography 40: 1278–1289.

Urbanitsa K, Landolt E. 1978. Recherches démographiques et écologiques sur une population hybridiognée de Cardamine L. Berichte des Geobotanischen Institutes der Eidgenössischen Technischen Hochschule, Stiftung Rübel Zürich 45: 30–53.

Wake DB, Haddy EA, Ackerly DD. 2009. Biogeography, changing climates, and niche evolution. Proceedings of the National Academy of Sciences, USA 106: 19631–19636.

Yamasaki E, Altermani F, Cavender-Bares J, Schuman MC, Zuppinger-Dingley D, Barona I, Schneider FD, Guillén-Escribá A, van Moorsel SJ, Halil Tet al. 2017. Genomics meets remote sensing in global change studies: monitoring and predicting phenology, evolution and biodiversity. Current Opinion in Environmental Sustainability 29: 177–186.

Yang C, L, Dong X, Cai Z, Tian W, Wang X. 2014. Short-term and continuing stresses differentially interplay with multiple hormones to regulate plant survival and growth. Molecular Plant 7: 841–855.

Yatsu Y, Kachi N, Kudoh H. 2003. Ecological distribution and phenology of an invasive species, Cardamine hirsuta L., and its native counterpart, Cardamine flexuosa With., in central Japan. Plant Species Biology 18: 35–42.

Yoo M-J, Szadkowski E, Wendel JF. 2013. Homeolog expression bias and expression level dominance in allopolyploid cotton. Heredity 110: 171–180.

Supporting Information

Additional Supporting Information may be found online in the Supporting Information section at the end of the article.

Dataset S1 List of genes whose homeolog expression ratio of Cardamine flexuosa differed between site-pairs on each sampling date and date-pairs at each sampling site.

Dataset S2 List of gene ontology (GO) terms for the biological process that were significantly enriched for each site-pair on each sampling date in Cardamine flexuosa.

Fig. S1 The distribution of the study species in Cardamine in Switzerland.

Fig. S2 Soil moisture of the Cardamine flexuosa sites where RNA samplings were conducted.

Fig. S3 The proportion of Cardamine hirsuta origin and C. amara origin reads in RNA-seq samples.

Fig. S4 Soil properties from 18 Cardamine sites in 2013.

Fig. S5 Sky openness from 18 sites measured at the season’s start (around Julian calendar day 80) and season’s end (around Julian calendar day 20) for Cardamine hirsuta sites and 170 for C. flexuosa and C. amara sites, respectively) in 2014.

Fig. S6 The result of principal component (PC) analysis on the expression level of 23 065 genes.

Fig. S7 The result of a clustering analysis based on log10FPKM calculated for 23 182 genes in Cardamine amara, C. hirsuta (H), C. amara homeolog of C. flexuosa (F3), and C. hirsuta homeolog of C. flexuosa (F3) on 2 May.
Fig. S8 Proportional Venn diagrams for the number of genes in the gene ontology (GO) category GO:0009414 (response to water deprivation) based on the GO enrichment analysis on the genes with homeolog expression ratio change in Cardamine flexuosa.

Fig. S9 The expression level of 32 genes in the gene ontology category GO:0009414 (response to water deprivation) for which the homeolog expression ratio of Cardamine flexuosa significantly differed between IR1 and KT5.

Methods S1 Detailed methods of data collection and analysis.

Table S1 The name of the study areas, the name and species composition (Type) of the study sites, and the number of individuals of the study species in Cardamine in 2013 and 2014.

Table S2 Pearson coefficients from pairwise correlations between environmental variables of Cardamine habitats.

Table S3 The developmental stage at sampling of the three individuals of Cardamine hirsuta, C. flexuosa, and C. amara in IR1, KT2, and KT5 in 2013.

Table S4 The numbers of reads mapped by HomeoRoq (H-origin: mapped to Cardamine hirsuta (2x); A-origin: mapped to C. amara (2x) and the number of expressed homeologs (FPKM > 1) in parenthesis shown for each individual of C. hirsuta, C. flexuosa (4x), and C. amara.

Table S5 The summary of the number of genes of Cardamine flexuosa examined and detected in analyses on site pairs.

Table S6 The summary of the number of genes of Cardamine flexuosa examined and detected in analyses on date pairs.

Table S7 Summary of gene ontology (GO) IDs and terms that were detected for the majority of site-pair by time combinations for genes whose homeolog expression ratio significantly changed in the allopolyploid Cardamine flexuosa (4x).

Table S8 List of Cardamine hirsuta genes and description of their Arabidopsis thaliana homologs belonging to the gene ontology (GO) ID GO:0009414 identified in the GO analysis on the homeolog expression ratio between at least one pair of sites on at least one date.

Please note: Wiley Blackwell are not responsible for the content or functionality of any Supporting Information supplied by the authors. Any queries (other than missing material) should be directed to the New Phytologist Central Office.