Received: 21 November 2018 | Revised: 21 June 2019 | Accepted: 3 July 2019

DOI: 10.1111/mec.15194

ORIGINAL ARTICLE

MOLECULAR ECOLOGY WILEY

Strengths and potential pitfalls of hay transfer for ecological restoration revealed by RAD-seq analysis in floodplain *Arabis* species

Hannes Dittberner[1] | Christian Becker[2] | Wen-Biao Jiao[3] | Korbinian Schneeberger[3] | Norbert Hölzel[4] | Aurélien Tellier[5] | Juliette de Meaux[1]

[1]Institute of Botany, University of Cologne, Cologne, Germany

[2]Cologne Center for Genomics, University of Cologne, Cologne, Germany

[3]Max-Planck-Institute for Plant Breeding Research, Cologne, Germany

[4]Institute of Landscape Ecology, University of Münster, Münster, Germany

[5]Center of Life and Food Sciences Weihenstephan, Technical University of Munich, Freising, Germany

Correspondence
Juliette de Meaux, Institute of Botany, University of Cologne, 50674 Cologne, Germany.
Email: jdemeaux@uni-koeln.de

Funding information
Deutsche Forschungsgemeinschaft, Grant/Award Number: 1529

Abstract

Achieving high intraspecific genetic diversity is a critical goal in ecological restoration as it increases the adaptive potential and long-term resilience of populations. Thus, we investigated genetic diversity within and between pristine sites in a fossil floodplain and compared it to sites restored by hay transfer between 1997 and 2014. RAD-seq genotyping revealed that the stenoecious floodplain species *Arabis nemorensis* is co-occurring with individuals that, based on ploidy, ITS-sequencing and morphology, probably belong to the close relative *Arabis sagittata*, which has a documented preference for dry calcareous grasslands but has not been reported in floodplain meadows. We show that hay transfer maintains genetic diversity for both species. Additionally, in *A. sagittata*, transfer from multiple genetically isolated pristine sites resulted in restored sites with increased diversity and admixed local genotypes. In *A. nemorensis*, transfer did not create novel admixture dynamics because genetic diversity between pristine sites was less differentiated. Thus, the effects of hay transfer on genetic diversity also depend on the genetic make-up of the donor communities of each species, especially when local material is mixed. Our results demonstrate the efficiency of hay transfer for habitat restoration and emphasize the importance of prerestoration characterization of microgeographic patterns of intraspecific diversity of the community to guarantee that restoration practices reach their goal, that is maximize the adaptive potential of the entire restored plant community. Overlooking these patterns may alter the balance between species in the community. Additionally, our comparison of summary statistics obtained from de novo- and reference-based RAD-seq pipelines shows that the genomic impact of restoration can be reliably monitored in species lacking prior genomic knowledge.

KEYWORDS

genetic diversity, hybridization, population structure, RAD-seq, reference genome, restoration genetics

1 | INTRODUCTION

Habitat degradation is an ever growing problem in our modern world, causing unprecedented loss of biodiversity and essential ecosystem services (Baillie, Hilton-Taylor, & Stuart, 2004). Thus, there is a growing demand for ecological restoration, that is measures assisting the recovery of ecosystems that have been degraded, damaged or destroyed (Society for Ecological Restoration International Science & Policy Working Group, 2004). However, ecological restoration is a difficult process rarely leading to full ecosystem recovery (Benayas, Newton, Diaz, & Bullock, 2009). Thus, the young field of restoration ecology, which studies ecological processes in the light of restoration, is vital to improve restoration practices (Bullock, Aronson, Newton, Pywell, & Rey-Benayas, 2011; Roberts, Stone, & Sugden, 2009; Suding, 2011).

One of the main goals of ecological restoration is the recovery of biodiversity, including both species richness and intraspecific genetic diversity (henceforth called genetic diversity). Genetic diversity has generally positive effects on ecosystems (reviewed in Hughes, Inouye, Johnson, Underwood, & Vellend, 2008). For example, experimentally increasing the genetic diversity of *Solidago altissima* increased primary above-ground biomass productivity and arthropod diversity (Crutsinger et al., 2006). Genetic diversity also boosts resistance of populations to invasion and environmental fluctuations, presumably because it enhances the adaptive potential of populations (Reed & Frankham, 2003; Vrijenhoek, 1994). For example, high diversity experimental *Arabidopsis thaliana* populations showed higher resistance against invasion by *Senecio vulgaris* than low diversity populations (Scheepens, Rauschkolb, Ziegler, Schroth, & Bossdorf, 2017). Moreover, *Zoestra marina* populations with higher genetic diversity showed increased biomass production, plant density and faunal abundance during an extremely warm period (Reusch, Ehlers, Hämmerli, & Worm, 2005). Concordantly, restored populations of *Z. marina* with increased genetic diversity showed longer plant survival, grew more rapidly and provided enhanced ecosystem services, measured by increased primary productivity, invertebrate density and nitrogen retention. This effect was stable in a range of environmental conditions along a water-depth gradient (Reynolds, McGlathery, & Waycott, 2012). These examples demonstrate the importance of genetic diversity for ecosystem function and stability and hence the need to consider population genetics in the design and evaluation of restoration practices, which is the focus of restoration genetics.

Restoration genetics can not only inform the planning of restoration efforts, for example, by identifying suitable source populations, but also help evaluating the success of restoration projects, for example, by monitoring genetic diversity in restored populations (Mijangos, Pacioni, Spencer, & Craig, 2015; Williams, Nevill, & Krauss, 2014). In fact, studies comparing the level of genetic diversity in pristine and restored populations frequently report limited success, with a reduction of genetic diversity in restored populations. This decline in genetic diversity may be caused by genetic bottlenecks in plant nurseries, biases introduced by seed harvesting strategies, founder effects during recolonization and/or unreliable commercial seeds (reviewed in Mijangos et al., 2015). By contrast, the transfer of seed-containing hay from pristine (donor) to restoration (donee) sites, termed hay transfer, is expected to limit the loss of genetic diversity and maintain site-specific local adaption (Hufford & Mazer, 2003; Kiehl, Kirmer, & Shaw, 2014). In addition, this method has the unique feature that it can, theoretically, restore an entire community without altering the genetic composition of populations and thus is the best method available for restoring entire ecosystems (Hölzel & Otte, 2003; Kiehl, Kirmer, Donath, Rasran, & Hölzel, 2010). So far, however, there is no empirical support for the efficiency of this practice (Bucharova et al., 2017), especially since many species maintain seed banks in the soil. Indeed, the genetic diversity specific to the seed bank will not be sampled with the hay, although it is known that it can contribute significantly to the maintenance of diversity (Tellier, Laurent, Lainer, Pavlidis, & Stephan, 2011).

The field of restoration genetics has witnessed a major technological shift over recent years. Restoration genetics studies have initially relied on microsatellites and AFLP markers and thus provided a limited overview on patterns of genetic variation within and between restored or pristine populations (reviewed in Mijangos et al., 2015). Now, genotyping-by-sequencing (GBS) methods are beginning to be more broadly adopted (Gruenthal et al., 2014; Massatti, Doherty, & Wood, 2018; O'Leary, Hollenbeck, Vega, Gold, & Portnoy, 2018; Torres-Martinez & Emery, 2016). These methods drastically reduce sequencing costs through strategies to sequence a reduced portion of the genome, for example, restriction site-associated DNA sequencing (RAD-seq; Elshire et al., 2011; Etter, Bassham, Hohenlohe, Johnson, & Cresko, 2011; Peterson, Weber, Kay, Fisher, & Hoekstra, 2012). In contrast to previous methods (AFLP, microsatellites), GBS approaches sample proportions of the genome that are sufficiently large to allow resolving patterns of genetic diversity and spatial structure even at very local scale where overall levels of genetic diversity are low (Bradbury et al., 2015; Jeffries et al., 2016; Reitzel, Herrera, Layden, Martindale, & Shank, 2013). In principle, GBS approaches have a third major advantage: they are well suited to unravel genetic diversity in nonmodel species without prior genomic information. Yet, the accuracy of genotyping in the absence of a reliable reference genome has been questioned (Shafer et al., 2016). Since target species in restoration projects rarely coincide with species or genera with advanced prior genomic knowledge, it is important to assess whenever possible, whether conclusions from RAD-seq-based restoration genetics studies depend on the availability of a reference genome.

Floodplain meadows are species-rich ecosystems, accommodating many endangered and stenoecious species adapted to the variable moisture regime. Due to increased agricultural land-use and river regulation, a large proportion of floodplain ecosystems are degraded and have become a target for ecological restoration. Here, we use RAD-seq to evaluate whether genetic diversity is maintained in floodplain meadows restored by hay transfer in the Upper Rhine valley in Germany (Donath, Bissels, Hölzel, & Otte, 2007; Hölzel & Otte, 2003). We focused on *Arabis nemorensis*, a stenoecious species

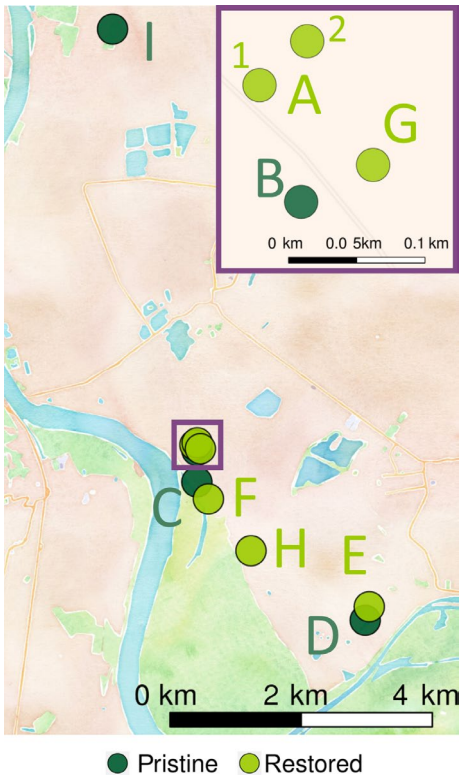

FIGURE 1 Overview map of studied sites. Each point on the map represents one site. The colour of the points represents the type of the site, that is whether it is pristine or restored. The purple square represents the area shown in the zoomed inset. All sites, except site F, are protected by a dyke

typically restricted to floodplain meadows, and thus strongly endangered in Central Europe (Schnittler & Günther, 1999). *Arabis nemorensis* is a short-lived, mostly biennial hemicryptophyte that is known to maintain a long-lived soil seed bank (Hölzel & Otte, 2004). It is part of the *Arabis hirsuta* species aggregate, which comprises several morphologically similar but ecologically diverse species, which diverged about 1.2 million years ago (Karl & Koch, 2014).

We performed RAD-seq for over 130 plants collected across pristine and restored sites. This data set allows us to ask the following questions: (a) What is the level of genetic diversity and structuration of the pristine sites that served as source populations for restoration? (b) Do restored sites show a lower level of genetic diversity than the pristine sites? (c) How did restoration affect the distribution of diversity within and among restored sites? (d) Is the use of a reference genome necessary to reliably characterize the impact of restoration on genetic diversity? This work demonstrates that a thorough genetic analysis of source and restored sites reveals the complex dynamics at stake in the restoration process. Through genetic analysis, we reveal the unanticipated co-occurrence of *A. nemorensis* with *A. sagittata*, a morphologically similar and closely related species from the *A. hirsuta* aggregate, which normally exhibits a preference for drier habitats such as calcareous grasslands (Hand & Gregor, 2006). We further show that restoration by hay transfer has maintained and can even enhance local diversity. Our analysis further shows that the use of a reference genome yields higher estimates of genetic diversity, but does not affect the resolution of patterns of genetic variation within and between sites, indicating that this approach can find broad applications in the field.

2 | METHODS

2.1 | Plant material and DNA extraction

The sampling area comprises the fossil, dyke-protected floodplain of the River Rhine near Riedstadt in Hessen, Germany. The area is dominated by arable fields, but also contains remnants of pristine flood meadow communities in low-lying depressions that are submerged by ascending groundwater during high floods of the River Rhine. Since ca. 20 years, new flood meadow communities have been restored on ex-arable land using the transfer of green hay from the pristine sites sampled in this study as donors to overcome significant dispersal limitation (Hölzel & Otte, 2003). During this process, hay from different donor sites was placed in distinct yet adjacent patches, making admixture possible. Since restoration is still ongoing, restored sites differ in age (Table S1).

From previous monitoring of the sites, we knew that populations of *Arabis* plants were not present every year in all sites. Thus, we sampled in two consecutive years to maximize the number of study sites. In the few sites sampled in both years, the genetic composition of the samples was not markedly different across years, so samples from both years were bulked for these sites (see Table S1 and Figure S1). We harvested seeds from a total of 134 plants of *Arabis nemorensis/sagittata* in nine sites named from A to I, in order of collection. The presence of *A. sagitata* was not previously reported in these sites and was thus unexpected. Although some individuals showed the reduced stem leaf density and shorter siliques typical of *A. sagittata*, these phenotypic criteria were not always clearly distinguishable in the field, especially at the end of the season when siliques matured, so that the presence of the two species has remained overlooked in previous studies (Burmeier, Eckstein, Donath, & Otte, 2011). Species identity was thus determined by post hoc analysis, after the RAD-seq analysis revealed the presence of two taxonomic units.

Four sites were sampled in pristine habitat (B, C, D and I) and five in restored habitat (A, E, F, G and H, Figure 1 and Table S2). In site A, two distinct stands of plants separated by about 100 m were sampled and were treated as subsites A-1 and A-2 throughout this study. To produce material for DNA extraction, we stratified seeds on wet filter paper for 6 days at 4°C in darkness. Afterwards, we sowed seeds in soil (33% VM, 33% ED-73, 33% Seramis (clay granules)). After 4–6 weeks of growth in the greenhouse, we harvested about 200 mg of leaf material from one offspring of each wild parent (genotype). We homogenized freshly harvested leaf material using a Precellys Evolution homogenizer (Bertin technologies) for 2 × 20 s at 6,800 rpm. We extracted DNA using the NucleoSpin Plant II Mini kit (Macherey-Nagel) following the manufacturer's instructions. We verified DNA quality using gel electrophoresis with a 0.8% agarose

gel. We measured DNA quantity using Qubit (broad-range kit) following manufacturer's instructions.

2.2 | Draft genome assembly and annotation

To facilitate genotype calling, we assembled a draft genome for one *A. nemorensis* accession (ID 29). Library preparation and sequencing was done at the Cologne Center for Genomics. Details about the sequencing, assembly and annotation of the genome are described in Appendix S1. Briefly, three libraries with insert sizes of 280 bp, 3 kbp and 6 kbp, respectively, were created and sequenced as 150 bp paired-end reads as part of an Illumina Hiseq 4000 lane for a total of 66 Gbp. We assembled reads using the *ALLPATHS-LG* assembler (Gnerre et al., 2011) with default settings. To further scaffold the genome, we generated 2.9 Gbp of PacBio sequence data. Library preparation and sequencing was done at the Max-Planck-Institute for Plant Breeding Research (Cologne, Germany). We scaffolded the genome using *OPERA-LG* with default settings (Gao, Bertrand, Chia, & Nagarajan, 2016). To achieve chromosome-level assembly, we created a whole-genome alignment with the *Arabis alpina* reference genome (Jiao et al., 2017; Willing et al., 2015) and performed synthetic path assembly using the tool *Synmap2* (Haug-Baltzell, Stephens, Davey, Scheidegger, & Lyons, 2017) available at the *CoGe* website (Lyons & Freeling, 2008). The pseudo-chromosomes had a total size of 192 Mbp and were used for all following analyses. The size of the genome was estimated by flow cytometry, which was performed commercially at Plant Cytometry Services (Didam, Netherlands). The estimated genome size was 274 Mbp. Thus, the assembly size was 70% of the genome size.

To detect and annotate transposable elements (TE), we used the softwares REPEATMODELER (Smit & Hubley, 2008) and REPEATMASKER (Smit, Hubley, & Green, 2013). Protein-coding genes were annotated by integrating predictions of ab initio gene annotation tools and alignments of homologous proteins. Three different tools including AUGUSTUS v3.2.3 (Stanke & Waack, 2003), GLIMMERHMM v3.0 (Majoros, Pertea, & Salzberg, 2004) and SNAP v2013 (Korf, 2004) were used to predict the initial gene models. Protein sequences from *A. thaliana*, *A. lyrata* and *A. alpina* (Arabidopsis Genome Initiative, 2000; Hu et al., 2011; Willing et al., 2015) were aligned to the assembly by the tool EXONERATE v2.2.0 (Slater & Birney, 2005). Then, the ab initio predictions and protein alignment hits were further combined to build the consensus gene models by the tool EVIDENCEMODELER (EVM) v2012 (Haas et al., 2008). Finally, TE-related genes in these models were annotated by checking the TE annotation, blastp (Altschul, Gish, Miller, Myers, & Lipman, 1990) alignments with Plant TE-related proteins and blastp alignments with *A. thaliana* proteins.

2.3 | RAD-sequencing and SNP calling

We genotyped 134 samples using the original RAD-sequencing (RAD-seq) protocol (Etter et al., 2011), with the following modifications (Appendix S1): (a) we used the enzyme KpnI-HF (New England Biolabs) for DNA digestion; (b) we ligated digested DNA with complementary adapters containing one of ten different barcodes and a stretch of five random nucleotides, used for post hoc removal of PCR duplicates (Table S3); (c) we created 14 pools of 10 barcoded samples each in equal amounts; and (d) we used indexed reverse primers for amplification, described in (Peterson et al., 2012), to allow multiplexing of pools. Libraries were sequenced on two Illumina HiSeq 4000 lanes with 2 × 150 bp.

We used FASTQC (Andrews, 2010) to quality-check the resulting reads. We trimmed adapters and removed reads shorter than 100 bp using CUTADAPT (Martin, 2011). We removed PCR duplicates based on a 5 bp stretch of random nucleotides at the end of the adapter, using the *clone_filter* module of STACKS version 1.37 (Catchen, Hohenlohe, Bassham, Amores, & Cresko, 2013). We de-multiplexed samples using the *process_radtags* module from STACKS. We filtered reads with ambiguous barcodes (allowed distance 2) and cut-sites, reads with uncalled bases and low-quality reads (default threshold).

For reference-based genotyping, we mapped reads using BWA (Li & Durbin, 2009) with default settings. We filtered mapped reads using SAMTOOLS (Li et al., 2009) and custom python scripts using the following criteria to remove reads: mapping quality <30, number soft-clipped bases >30; reads were unpaired; mates mapped on different chromosomes; mate mapping distance >700. The proportion of mapped reads varied slightly among species, with a mean of 48% for *A. nemorensis*, 43% for hybrids and 42% for *A. sagittata* (Table S4). We called genotypes using SAMTOOLS MPILEUP and VARSCAN2 (Koboldt et al., 2012) with the following options: base quality >20; re-calculation of base quality on the fly (-E option); read depth > 14; strand filter de-activated; SNP calling *p*-value < .01. We filtered genotyped loci using VCFTOOLS (Danecek et al., 2011) and custom python scripts removing loci with missing data in more than 5% of individuals and loci in masked (repetitive) regions. We clustered contiguous loci spaced less than 100 bp apart into RAD regions. Regions with excessively low or high coverage are likely results of allele dropout or paralogous mapping, respectively. Thus, we removed regions fulfilling one of the following criteria: mean coverage of the region greater than twice the overall mean coverage; mean coverage of the region smaller than a third of the overall mean coverage; region maximum coverage greater than twice the mean maximal coverage over all regions; region shorter than 250 bp; region longer than 1,300 bp.

After depth filtering, we still found loci with a high frequency of heterozygotes (up to 100%), which is unlikely in highly selfing species. Notably, in over 90% of these loci only one of the two homozygous genotypes was observed. Moreover, these highly heterozygous loci clustered in high density on single RAD fragments. Thus, these loci likely resulted from paralogous mapping, which passed the depth filter due to sequencing depth variation among RAD regions. Therefore, we removed RAD regions containing loci with a frequency of heterozygotes greater than 20%. This threshold was picked to limit the impact of noise from mapping artefacts while still allowing for reasonable levels of heterozygosity, since some low-level outcrossing likely occurs. From the resulting genotype data set, we extracted single nucleotide polymorphisms (SNPs) using VCFTOOLS (Danecek et al., 2011).

For de novo genotyping, we used the *Stacks* 2.2 *denovo_map.pl* pipeline (Catchen et al., 2013). The aim of the de novo analysis was to test whether one would reach similar results and conclusions without the use of a reference genome. Thus, we performed this analysis without knowledge about the two species (since we gained this from the reference-based analysis) and ran *Stacks* for all samples combined. As recommended by the authors of the tool (Rochette & Catchen, 2017), we first used a subset of 15 representative genotypes to tune the parameters (−*M* and −*n*) of the algorithm, which control the number of mismatches between stacks within (*M*) and between (*n*) individuals. We varied *M* and *n* from 1 to 9. For each set of parameters, we analysed the number of loci shared between 80% of the samples. This measure peaked at the value six for *M* and *n*. Thus, we used this value for both parameters for the full analysis. We ran the *denovo_map.pl* pipeline using .01 as the *p*-value threshold for calling genotypes and SNPs, and otherwise default options. We used the *populations* program to create a VCF file for further analysis using the following filters: 5% maximum missing data per locus; 20% maximum observed heterozygosity per locus; locus must be present in all sites.

2.4 | Population genetics statistics

We did all statistical analysis using R version 3.4.4 (R Development Core Team, 2008) and provide a supplemental R Markdown file (Appendix S2). The following packages were used for plotting: GGPLOT2 (Wickham, 2009), GGMAP (Kahle & Wickham, 2013), GGTHEMES (Arnold et al., 2017), GGSN (Baquero, 2017) and HEATMAP3 (Zhao, Guo, Sheng, & Shyr, 2015). We performed all analysis for the reference-based and the de novo-based data set and compared the results. We used the VCFR package (Knaus & Grünwald, 2017) to load VCF files into R and make the SNP data available for processing with other libraries. Based on our annotation, we determined whether SNPs are in coding regions and whether they are synonymous or non-synonymous using the POPGENOME package (Pfeifer, Wittelsbuerger, Li, & Handsaker, 2018). We performed principal component analysis (PCA) of SNP data for all samples using the ADEGENET package (Jombart et al., 2016). Missing data were scaled to the mean for PCA.

Based on the first principal component, most individuals (83%) could be assigned to one of two distinct taxonomic groups. We used molecular methods to assign species labels to the two taxonomic groups. We sequenced the internal transcribed spacer (ITS) sequence of nine individuals, three from the first (left) and five from the second (right) cluster and one located between the clusters (Table S5). Primers for amplification were taken from Mummenhoff, Franzke, and Koch (1997). We used the sequences as input to the taxonomy tool of the *Brassibase* website (Kiefer et al., 2014). As a complementary approach, we downloaded the 612 bp long reference ITS sequences of all species in the *Arabis hirsuta* group from Brassibase (Kiefer et al., 2014). Using the software SEAVIEW (Gouy, Guindon, & Gascuel, 2010), we aligned these sequences to the ITS sequences of our samples using the MUSCLE algorithm (Edgar, 2004; Appendix S3) and created a phylogeny using PHYML (Guindon & Gascuel, 2003) with default settings. The ITS sequence does not allow distinction between *A. sagittata* and *A. hirsuta*. Yet, the species can be distinguished based on ploidy as *A. sagittata* is diploid and *A. hirsuta* tetraploid (Karl & Koch, 2014). Leaf samples were therefore collected from 12 individuals (seven putative *A. sagittata*/*A. hirsuta*, two putative *A. nemorensis*, three putative hybrids; Table S6) and their relative DNA content per cell (in comparison with a company-internal standard [*Vinca minor*]) determined by Plant Cytometry Services (Didam, The Netherlands). Additionally, we included one sample from each of two independent *Arabis hirsuta* populations sampled in 2016 in Bavaria (see Table S2). We inferred the ploidy of our samples by comparing relative DNA content: individuals of the same ploidy should have similar DNA content, and tetraploids should have twice the DNA content of diploids.

We conducted all population genetic analyses separately for the two species. Individuals, which were not assigned to any species, were likely interspecific hybrids and excluded from further analysis. We used the *pegas* library to calculate within-site genetic diversity (Nei's π; average pairwise nucleotide differences) for each site (Paradis, Jombart, Schliep, Potts, & Winter, 2016), excluding sites with less than two individuals per respective species. To scale the estimates of π, we divided the average number of pairwise nucleotide differences among samples by the total number of successfully genotyped bases, excluding all bases which failed any of our previously described genotype filters (missing data, region heterozygosity, region depth and length). For the de novo pipeline, we extracted the total number of genotyped bases from *Stacks* output. We calculated correlation coefficients between reference-based and de novo-based π estimates using Pearson's method. We calculated pairwise F_{ST} (Nei, 1987) and genetic distance (Cavalli-Sforza & Edwards, 1967) between all pairs of sites using the HIERFSTAT package (Goudet & Jombart, 2015). Negative F_{ST} values were set to zero. Differences of genetic distance and F_{ST} among pristine and restored sites were tested using a Wilcoxon rank-sum test on pairwise distance matrices. We tested for correlation between the distance matrices of the reference and de novo data sets using a Mantel test with 10,000 permutations.

2.5 | Admixture analysis

For ADMIXTURE analysis (Alexander, Novembre, & Lange, 2009), we converted VCF files to bed-files using PLINK (Purcell, 2009; Purcell et al., 2007). First, we conducted ADMIXTURE analysis for all samples combined for $K = 1$ to $K = 10$ (reference-pipeline only). Then, for each of the species and pipeline (reference/de novo), we ran ADMIXTURE analysis for $K = 2$ to $K = 10$, with 10 iterations of cross-validation each. Before plotting, we normalized clusters across runs using CLUMPAK (Kopelman, Mayzel, Jakobsson, Rosenberg, & Mayrose, 2015). We created plots using a custom R-script.

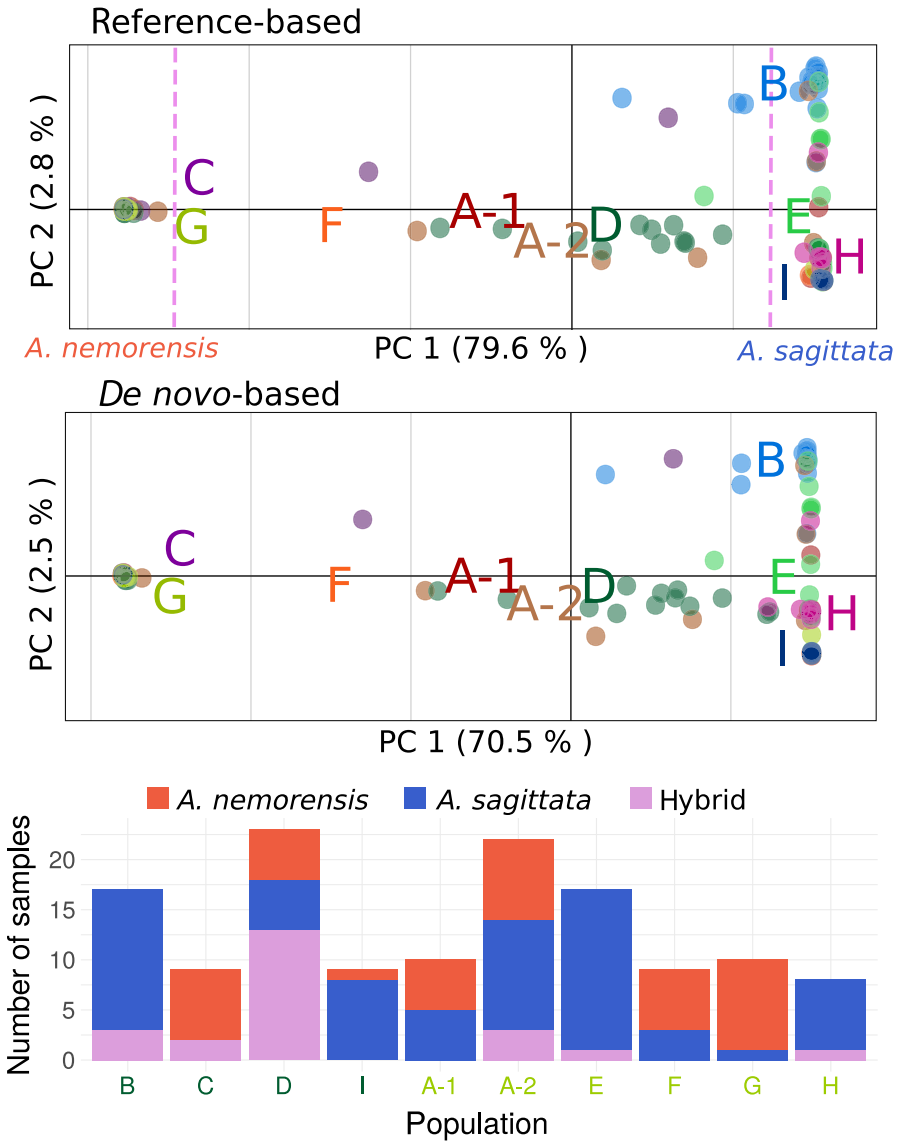

FIGURE 2 Presence of two hybridizing *Arabis* species in study sites. (Top) Plot of the first two principal components (PC) of genetic variation based on the reference-based pipeline. Colours distinguish the different sites, which are labelled with letters of the corresponding colour. Site labels are in the centroid of the respective site (with offset to avoid overlaps). Values in brackets in the axis labels show the amount of variance explained by each PC. Note that the first PC explains about 80% of the total variation and splits most individuals into two clearly defined clusters. The two clusters along PC1 correspond to two distinct species, identified here as *Arabis nemorensis (left)* and most probably *Arabis sagittata (right)*. Species identity is indicated by text labels below the plot. Points in between the two clusters are likely natural interspecific hybrids. The dashed lines show the thresholds used for distinction between species and hybrids throughout this study. (Middle) Plot showing the same analysis as above but based on the de novo pipeline. (Bottom) Stacked barplot showing the distribution of species among the study sites. Population labels are coloured by type: dark green = pristine, light green = restored. Species are distributed heterogeneously among sites

3 | RESULTS

3.1 | RAD-sequencing uncovers two hybridizing species

Unless otherwise stated, all described results were obtained using the reference-based pipeline. We genotyped 134 individuals from 10 sites—4 pristine and 6 restored (Figure 1)—yielding 3.6 Mb of sequence of which 32,880 single nucleotide positions were polymorphic (SNPs). Only 20% of SNPs were in coding regions, 40% and 56% of which were synonymous and nonsynonymous, respectively (4% unassigned). To visualize patterns of genetic diversity across sites, we conducted principal component analysis (PCA) for all individuals (Figure 2, top). Almost 80% of the total genetic variation was explained by the first principal component, which separated most individuals into two clearly defined clusters, likely representing taxonomic units. To determine species identity, we performed phylogenetic analysis of the ITS region of nine individuals (Table S5) using two different methods (see methods), which gave the same results. We confirmed that one taxonomic unit was *A. nemorensis* (Figures S2 and S3). The ITS region of individuals from the other taxonomic unit was identical to that of *A. sagittata* and *A. hirsuta*, a sequence that differs by at least three nucleotides from all other ITS sequences of known species of the complex (Figures S3 and S4). Yet, these sibling species can be distinguished based on ploidy, as *A. hirsuta* is tetraploid and *A. sagittata* diploid (Karl & Koch, 2014). Since all twelve tested samples from the Rhine populations had the same genome size, which was half of that of the two *A. hirsuta* samples (Table S6), we conclude that the second cluster most likely corresponds to the diploid species *A. sagittata*. Average genetic distance (d_{XY}) between *A. nemorensis* and *A. sagittata* was 8.1e−03, that is 8 fixed SNP differences for 1,000 bp.

Twenty-three individuals showed a positioning along the first PC that was intermediate between the two clusters, suggesting they were interspecific hybrids. Most of these hybrids were closer to *A. sagittata* on the first PC. Since F1 hybrids should be located

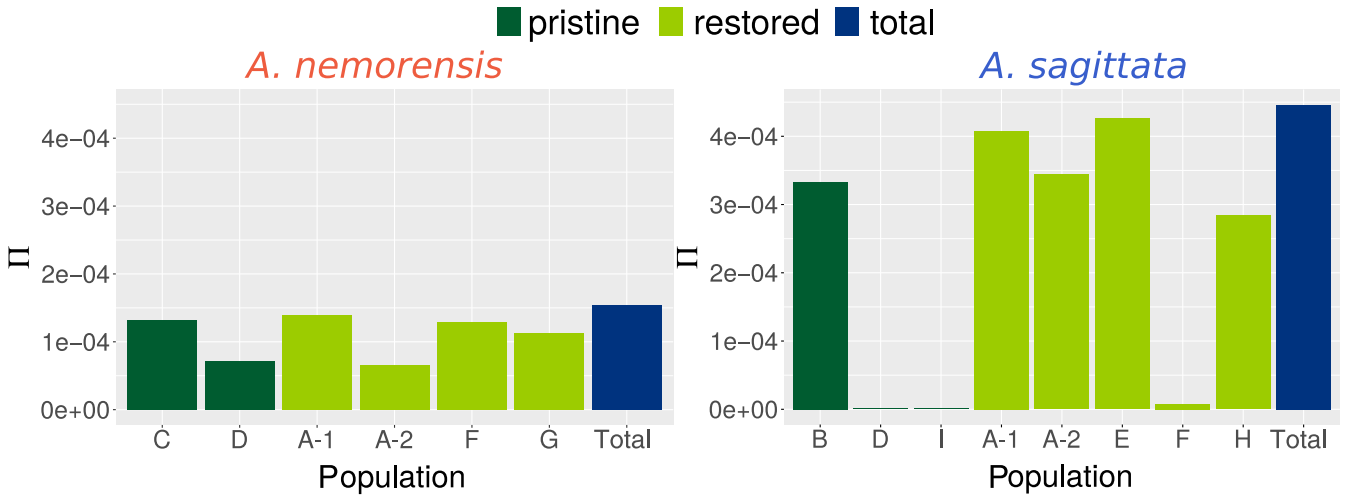

FIGURE 3 No reduction in genetic diversity in restored sites. Barplot of average pairwise genetic diversity (π) within each site of each species. Bar colour indicates the type of site

exactly in the middle between the two species, this suggests that they are somewhat fertile and preferentially back-cross with *A. sagittata*. Laboratory observations have shown that hybrids are indeed fertile (H. Dittberner, personal observation). Additionally, in an ADMIXTURE analysis of all samples with $K = 2$, most hybrids showed ancestry over 50% from *A. sagittata*, as expected from preferential back-crossing to this parent (Figure S5).

Overall, our sample was composed of 31% *A. nemorensis*, 52% *A. sagittata* and 17% hybrids. The species composition differed among sites (Figure 2, bottom). *A. nemorensis* was present in 7 sites (3 pristine, 4 restored), *A. sagittata* in 9 sites (3 pristine, 6 restored) and hybrids in 6 sites (3 pristine, 3 restored). Notably, the pristine site D was dominated by hybrids with over 56%.

3.2 | No reduction of genetic diversity in restored sites

We computed species-specific estimates of genetic diversity within each site, excluding hybrid genotypes. The *A. nemorensis* data set consisted of 2,746 SNPs. Levels of genetic diversity (π) varied up to twofold among sites, ranging from 6.6e−05 in A-2 to 1.4e−04 in A-1 (Figure 3, left; Table S7). Total diversity was 1.5e−04. However, pristine and restored sites did not differ significantly in their level of diversity (mean difference = +10% in restored; $W = 4$, $p = 1$).

The *A. sagittata* data set consisted of 6,366 SNPs. Total genetic diversity in *A. sagittata* was about three times as high as in *A. nemorensis*. Yet, in contrast to *A. nemorensis*, genetic diversity differed strongly among sites, ranging from 1.03e to 06 in site I to 4.26e−04 in site E (Figure 3, right; Table S7). Notably, genetic diversity was low in two of three pristine sites. In contrast, we found high levels of diversity in all restored sites, except site F. However, the overall difference between pristine and restored sites was not significant (mean difference = +163% in restored sites; $W = 13$, $p = .14$).

Since hybridization potentially enables gene flow between the two species, we also compared levels of genetic diversity for both species combined, including the hybrids. Overall, genetic diversity increased by an order of magnitude and mixed sites were more diverse than mostly pure sites, as would be expected (Figure S6). Again, restored sites did not show significantly different levels of genetic diversity from pristine sites (mean difference = +22%; $W = 13$, $p = .91$).

3.3 | Restoration reduces population structure and facilitates recombination

To quantify the degree of population structure, we estimated genetic distance and differentiation (F_{ST}) among all pairs of sites (Table S8). Genetic distance among *A. nemorensis* sites ranged from 0.03 to 0.31 and F_{ST} estimates from 0 to 0.5 (Figure 4, A + C). Population structure was slightly more pronounced among pristine sites than restored sites: mean genetic distance was 0.26 among pristine sites and 0.13 among restored sites ($W = 17$, $p = .047$); mean F_{ST} was 0.37 among pristine sites and 0.25 among restored sites ($W = 14$, $p = .26$).

In *A. sagittata*, genetic distance ranged from 0.32 to 0.34 and F_{ST} estimates from 0 to .91 (Figure 4, B + D). The difference in population structure between pristine and restored sites was stronger than in *A. nemorensis*: mean genetic distance was 0.33 among pristine sites and 0.12 among restored sites ($W = 45$, $p = .002$); mean F_{ST} was 0.81 among pristine sites and 0.2 among restored sites ($W = 43$, $p = .015$).

Reduced differentiation among restored sites suggests that genetic material of pristine sites was mixed in restored sites. To visualize this in more detail, we conducted ADMIXTURE analysis for each species. This analysis assumes a given number (K) of genetic clusters (ancestral populations) and assigns cluster ancestry proportions to each individual, allowing for mixed ancestry. For *A. sagittata*, $K = 6$ was estimated as the optimal value (Figure 5). Up to $K = 4$, the three pristine sites consisted of three distinct genetic clusters, which were mixed in all but one of the restored sites. For higher values of K, pristine population B consisted of two clusters, with pure and admixed individuals in equal proportions, in agreement

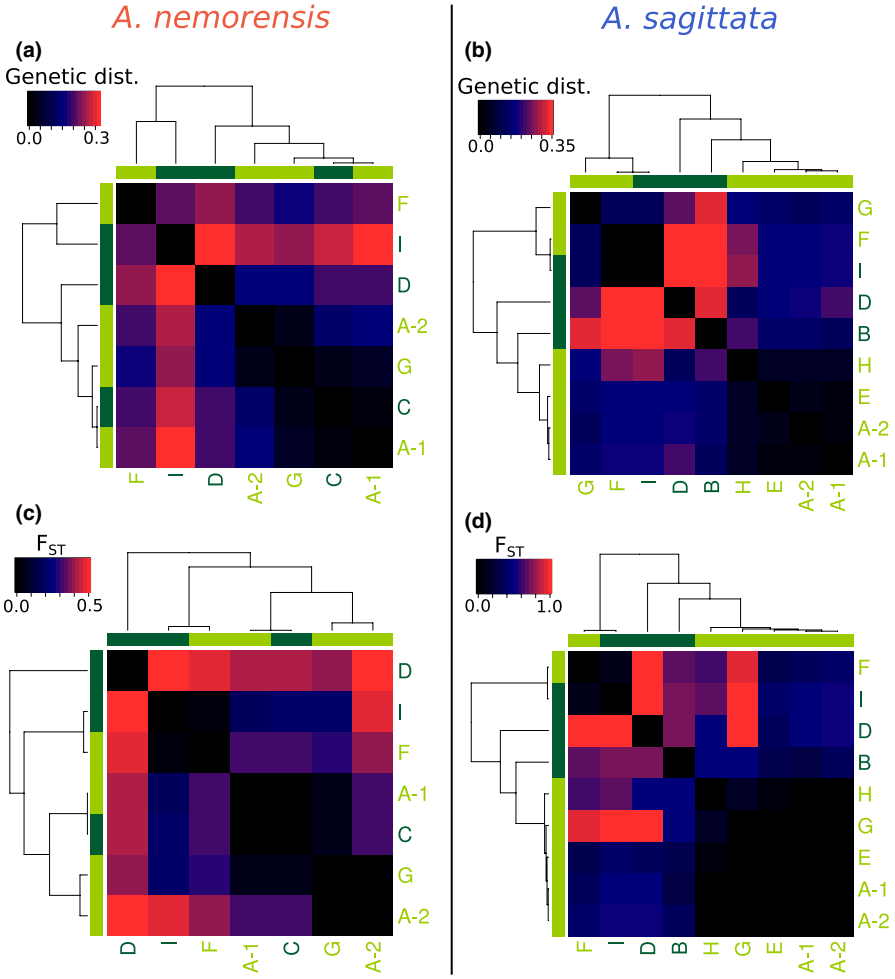

FIGURE 4 Overall strong population structure with signatures of admixture in restored sites. This panel shows different measurements of population structure of each species. (A–D) Heatmaps showing pairwise comparisons between all sites of genetic distance (Cavalli-Sforza chord distance) and F_{ST}, respectively. Values of the respective variable are indicated by tile colour, corresponding to the colour scale in the top-left corner of each heatmap. Sites were clustered based on the respective variable, as indicated by the dendrograms. Colours of text labels and bars next to the grid indicate the type of the site (same colour scheme as Figure 3)

with the higher genetic diversity observed at this site. All individuals from other pristine sites had pure ancestry from a single cluster per population.

For *A. nemorensis*, $K = 7$ was estimated as the optimal value for clustering (Figure 5). In contrast to *A. sagittata*, we found individuals with different or mixed ancestry within populations of *A. nemorensis*, even at low values of K, indicating more genetic mixture than in *A. sagittata*. Since genetic structure was less pronounced for *A. nemorensis*, the restoration procedure did not impact the genetic distribution of this species, by contrast with *A. sagittata*.

3.4 | De novo- and reference-based summary statistics reach the same conclusion

Finally, we tested whether a reference genome is required to determine the impact of restoration on genetic diversity by comparing results from a reference-based and a de novo pipeline. We found that estimates of genetic diversity, genetic distance or F_{ST} yielded by the two methods were highly correlated, with all correlation coefficients being greater than 0.95 (maximum $p < .001$, Figure 6). However, we observed that estimates of genetic diversity generated without a reference genome were deflated, especially for high diversity sites (Figure 6), by a median factor of 0.83 for *A. nemorensis* and 0.36 for *A. sagittata*. Estimates of genetic distance were underestimated by a median factor of 0.61 in both species. Interestingly, however, both pipelines yielded almost identical estimates of F_{ST} (median factor of 0.93) and revealed a very similar extent of admixture between sites (Figures S7 and S8). Moreover, the presence of the two species and their hybrids was detected with both methods (Figure 2 middle). We concluded that, in this study, the use of a reference genome was not required to determine the impact of restoration on genetic diversity. Since the two pipelines coincide, we could further conclude that the distribution of variation in two species and across pristine and restored sites reported above is not the result of possible mapping biases to the reference genome.

4 | DISCUSSION

4.1 | A reference genome is not required to characterize the impact of restoration

While RAD-seq and related methods are a cheap tool to acquire genotype information across the genome without need for a reference genome (Elshire et al., 2011; Etter et al., 2011; Peterson et al., 2012), the reliability of de novo assembly pipelines has been questioned (Shafer et al., 2016). The availability of a reference genome allowed

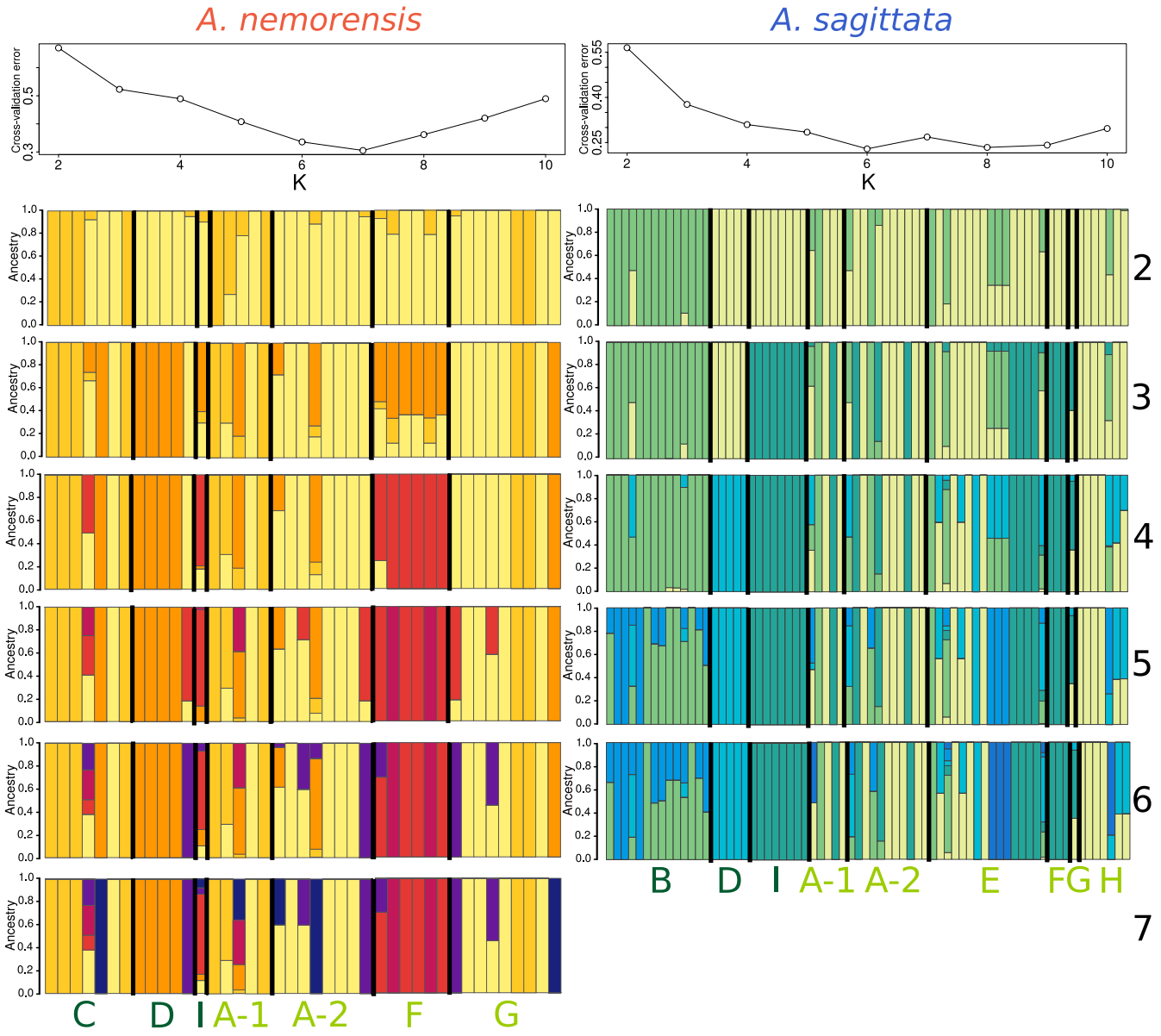

FIGURE 5 Increased impact of genetic admixture of pristine sites in *A. sagittata*. The figures shows a comparison of ADMIXTURE results for both species. The top panels show the cross-validation (CV) error for all tested values of *K*. The following panels show the ancestry proportions of each cluster for each individual for all values of *K* until the first minimum of the CV-error curve. The numbers on the right represent the value of *K*. Site labels are coloured according to the type of the site: light green = restored, dark green = pristine

us to ask whether reference-based read mapping pipelines yielded distinct conclusions from a pipeline based on de novo read assembly. Conversely, this also allowed testing whether the use of an *A. nemorensis* reference genome to map *A. sagittata* samples could have biased our conclusions. The results from both pipelines were highly correlated. Thus, comparative analysis of sites was reliable with both pipelines. However, the de novo pipeline (*Stacks*, Catchen et al., 2013) underestimated the amount of genetic diversity compared to the reference-based pipeline. This is in contrast to a previous study comparing different RAD-seq pipelines (Shafer et al., 2016), where *Stacks* produced slightly inflated estimates compared to reference-based pipelines. However, in the same study, other de novo pipelines produced substantially lower estimates of genetic diversity (Shafer et al., 2016). Thus, the magnitude of genetic diversity estimates might depend on the study system and pipelines/parameters used, and caution is advised when comparing these estimates across studies.

In contrast, both pipelines agreed for analyses comparing diversity between species or sites (e.g., F_{ST}, ADMIXTURE). Thus, we conclude that RAD-seq is an efficient tool to characterize the distribution of diversity, even in the absence of a reference genome. We therefore hope that this study will pave the way for exploring how species, with diverse life history and uncharacterized genomes, will be maintained after hay transfer or how modalities of hay transfer affects not only single species but the balance between multiple species in the community.

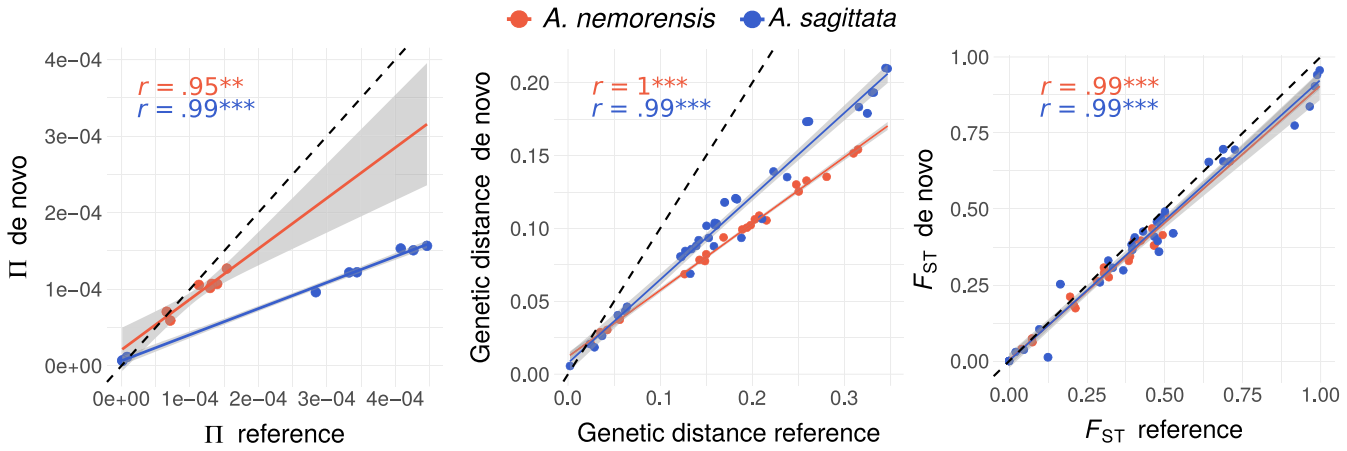

FIGURE 6 De novo- and reference-based summary statistics were highly correlated. Plots show the correlation of summary statistics based on the reference-based and the de novo-based pipeline. Line and dot colours represent the species. Each dot represents one individual. Lines represent a linear fit through the dots, and the grey shadows indicate the error of the fit. The dashed line represents a hypothetical 1:1 relationship of the variables. Correlation coefficients (Pearson's product-moment correlation) are indicated as coloured text corresponding to the species. Stars indicate the level of the coefficient: $^{**}p < .01$, $^{***}p < .001$

4.2 | First documented presence of *A. sagittata* in floodplain meadow habitat

We detected two independent genetic clusters. Based on ITS sequences, ecological preference and morphological descriptors, the first one appears to represent *bona fide A. nemorensis* individuals. Taxonomic boundaries in the *Arabis hirsuta* clade can be tenuous, yet ITS sequence differences between the two clusters strongly suggest that they belong to distinct species (Karl & Koch, 2014). ITS sequences, ploidy level and morphological evidence collectively indicated that individuals in the second genetic cluster are likely to belong to *Arabis sagittata*. Species barriers are often incomplete, and patterns of divergence between closely related species of the same genus can be strongly intertwined. Population genomics approaches and wider sampling will be needed to consolidate our understanding of the *A. sagittata* taxonomic group and its historical relationship with the *A. nemorensis* taxon. It is indeed possible that gene flow between the species is ancient and not restricted to the Rhine population. Population genomics can indeed yield detailed view on the process and history of speciation in plant genera (see, e.g., Novikova et al., 2016).

We did not anticipate the presence of the *A. sagittata* in floodplain habitats because it is predominantly found in warm and dry habitats (Hand & Gregor, 2006). However, agricultural land-use and flood regulation have considerably modified the floodplain ecosystem: flooding is contained by a dyke and the ground water level has decreased (Hölzel & Otte, 2001). Artificial modifications of the environment can change selection regimes and facilitate the establishment of non-native species (Byers, 2002; Crooks, Chang, & Ruiz, 2011; Fukasawa, Miyashita, Hashimoto, Tatara, & Abe, 2013; Tyrrell & Byers, 2007). Thus, it is possible that *A. sagittata* migrated into the floodplain ecosystem after human river regulation decreased the frequency and severity of flooding. We want to stress that the species co-occurred both in pristine and restored sites. Thus, the contact between the species was not caused by the restoration efforts. In fact, we observed that *A. sagittata* is more frequent than *A. nemorensis* in our sample, which raises the concern that this species may be in the process of displacing *A. nemorensis*. This trend could be enhanced by climate change, which may lead to conditions that further favour the xero-thermophylic *A. sagittata*. Additionally, if *A. sagittata* is often mistaken for *A. nemorensis* in flora reports on floodplain environments, the remaining *A. nemorensis* population in Central Europe might be even smaller than is currently assumed.

4.3 | The parallel transfer of two sympatric relatives shows that hay transfer maintains genetic diversity

Hay transfer has proven to be a particularly successful method for establishing new populations of target species in ecological restoration across a variety of herbaceous vegetation types (Coiffait-Gombault, Buisson, & Dutoit, 2011; Hölzel & Otte, 2003; Kiehl et al., 2010; Kiehl & Wagner, 2006; Török et al., 2012). Furthermore, hay transfer is often seen as the gold standard to preserve local levels of genetic diversity and adaptation (Vander Mijnsbrugge, Bischoff, & Smith, 2010). The latter is probably the main reason, why it is increasingly used in ecological restoration (Kiehl et al., 2014). The aim of this study was to characterize the level of diversity in the pristine source sites and document the impact of hay transfer on the genetic diversity in restored sites. Although our initial plan was to focus on *A. nemorensis*, a typical representative of species-rich floodplain meadows, the unanticipated presence of *A. sagittata* in our sample allowed us to compare the genetic effects of restoration by hay transfer on the two species.

Several years after restoration, we did not find a significant difference in genetic diversity between pristine and restored sites for either of the two species. Thus, the hay transfer method can restore populations with levels of diversity indistinguishable from the source populations in the long term. Our findings are in agreement

with studies on population life-stage structure and dynamics comparing pristine and restored sites of *A. nemorensis*/*A. sagittata* in the same region (Burmeier et al., 2011). This outcome is particularly remarkable given that *A. nemorensis* has a long-term seed bank, with up to 25,000 germinable seeds*m^{-2}, which was not transferred to restored sites (Burmeier et al., 2011). Although species for which the genetic diversity present in the seed bank tends to differ more strongly from the above-ground diversity may fare differently after hay transfer, we note that populations restored with alternative methods, for example spontaneous recolonization (Vandepitte et al., 2012) or propagated seed mixtures (Espeland et al., 2017; Fant, Holmstrom, Sirkin, Etterson, & Masi, 2008), both excluded the seed bank and revealed a reduction of diversity (Mijangos et al., 2015). Thus, hay transfer might be superior to other restoration methods not only in restoration success (Hölzel & Otte, 2003; Kiehl et al., 2010) but also in transferring genetic diversity.

4.4 | The modalities of habitat restoration can modify the relative adaptive potential of species in the ecosystem

Genetic diversity was low in *A. nemorensis* and *A. sagittata*, compared to other outcrossing or selfing Brassicaceae species (Mattila, Tyrmi, Pyhäjärvi, & Savolainen, 2017; Onge, Källman, Slotte, Lascoux, & Palmé, 2011). In fact, genetic distance between species was only on the order of magnitude as diversity found within populations of *Arabidopsis lyrata* (Mattila et al., 2017). Genetic diversity within our species was similar to that reported for *Arabis alpina* populations in Scandinavia (Laenen et al., 2018). As in our study, these populations are selfing and located on the margin of the species' range (Jalas & Suominen, 1994), two factors often coinciding with lower levels of genetic diversity. Since populations with low genetic diversity may suffer from increased genetic load and decreased adaptive potential, the transfer of nonlocal material is often envisaged to preserve endangered species (Breed, Stead, Ottewell, Gardner, & Lowe, 2013; Weeks et al., 2011). This approach is however controversial. Strategies that introduce nonlocal seeds can decrease population fitness either by introducing maladapted genotypes (Crémieux, Bischoff, Müller-Schärer, & Steinger, 2010; McKay, Christian, Harrison, & Rice, 2005) or by causing outbreeding depression (Frankham et al., 2011). As a compromise, a strategy of mixing regional seeds was recently proposed (Bucharova et al., 2018). Our results show that this strategy was unintentionally implemented in the examined restoration effort: ADMIXTURE analysis showed that pristine *A. sagittata* sites are dominated by one or two ancestral groups, with very low genetic diversity within each group. Yet, in some of the restored sites, admixture of low diversity *A. sagittata* groups took place, leading to a strong increase in genetic diversity and decrease in population structure. ADMIXTURE analysis also revealed that this led to genetic recombination between distinct haplotypes in *A. sagittata*, possibly increasing the adaptive potential of restored sites in this species. Increased genetic connectivity between populations has indeed been shown to help maintain or even increase genetic diversity in the long term (DiLeo, Rico, Boehmer, & Wagner, 2017). Fitness assays of recombined *A. sagittata* genotypes are needed to verify whether this local admixture has reinforced the establishment of this species in restored floodplain meadows.

Interestingly, we find less pronounced population structure in pristine sites of *A. nemorensis* than in *A. sagittata*, suggesting a comparatively higher level of gene flow. The restoration of this species is therefore less likely to benefit from post-transfer admixture. More so, our results suggest that it is possible that *A. nemorensis* is at increased disadvantage in restored habitats, if admixture and resulting recombination were to favour the emergence of more competitive genotypes in *A. sagittata*.

The coexistence of these species in the Rhine floodplain ecosystem will be further impacted by the ongoing hybridization dynamic that our analysis uncovers for the first time in this system and which occurred independently of the restoration effort. *A. nemorensis* could receive alleles conferring drought-adaptation from the xero-thermophilic *A. sagittata* that may enhance its ability to cope with increased drought exposure in its habitat. The genomic composition of the hybrids, however, also indicates that hybrids backcross preferentially with *A. sagittata*. Gene flow from *A. nemorensis* could facilitate adaptation of *A. sagittata* to the floodplain environment. Such adaptive introgressions could also potentially accelerate the extinction of *A. nemorensis*. While hybridization is common in plants (Mallet, Besansky, & Hahn, 2016), well-documented cases of adaptive introgression are rare and require elaborate experiments (Goulet, Roda, & Hopkins, 2017; Suarez-Gonzalez, Lexer, & Cronk, 2018). Such experiments are now warranted to determine the impact of hybridization in this sympatric species complex.

5 | CONCLUSIONS

Clearly, genetic analysis helps with species identification as sibling species can be difficult to distinguish morphologically even for specialists. A unique feature of hay transfer restoration approaches is that the whole plant community can be transplanted. It is therefore particularly important to determine the composition of the source populations to limit the spread of nontarget species (Bickford et al., 2007). Our study demonstrates that hay transfer has maintained genetic diversity in restored populations. However, we also note that it may have inadvertently contributed to increase the genetic diversity and adaptive potential of only one of the two species, due to differences in the genetic make-up of the donor populations. This might lead to a competitive advantage of one over the other species in the long term, potentially disturbing the balance in the community. On the one hand, this shows that restoration by hay transfer may enhance the adaptive potential, which is especially important in the face of a rapid climate change. On the other hand, this also highlights that understanding the underlying genetics of the community to be transferred is a prerequisite for the design of restoration strategies that promote the maintenance of both endangered ecosystems and endangered species.

ACKNOWLEDGEMENTS

We thank Matthias Harnisch for providing information about the populations, Markus Koch for insightful discussions about the *A. hirsuta* tribe, Eric Schranz for his assistance with the pseudo-chromosome assembly and Gregor Schmitz for helpful feedback. Further, we thank Janine Altmüller and the team of the Cologne Center for Genomics for their assistance in RAD-seq optimization, library preparation and sequencing. This work was partly funded by the German Research Foundation Deutsche Forschungsgemeinschaft DFG [DFG priority program 1529 'ADAPTOMICS'].

AUTHOR CONTRIBUTIONS

J.d.M., H.D., A.T. and N.H. conceived the study. H.D. and N.H. collected plant material; H.D. and C.B. prepared material for sequencing; W.B.J., K.S. and H.D. were responsible for bioinformatic processing. J.d.M., H.D. and N.H. analysed data and wrote the manuscript with significant contributions from C.B., W.B.J., K.S. and A.T.

DATA AVAILABILITY STATEMENT

The genome assembly, annotation, raw reads and RAD-seq samples were uploaded to EMBL ENA (ID: PRJEB33482). ITS sequences were uploaded to NCBI GenBank (MN166461–MN166469). VCF files and custom scripts were stored in a Dryad repository (DOI: https://doi.org/10.5061/dryad.t60vh3p).

ORCID

Hannes Dittberner https://orcid.org/0000-0002-4226-7606

REFERENCES

Alexander, D. H., Novembre, J., & Lange, K. (2009). Fast model-based estimation of ancestry in unrelated individuals. *Genome Research*, *19*(9), 1655–1664. https://doi.org/10.1101/gr.094052.109

Altschul, S. F., Gish, W., Miller, W., Myers, E. W., & Lipman, D. J. (1990). Basic local alignment search tool. *Journal of Molecular Biology*, *215*(3), 403–410. https://doi.org/10.1016/S0022-2836(05)80360-2

Andrews, S. (2010). FASTQC: *a quality control tool for high throughput sequence data (Version 0.11.3)*. Retrieved from http://www.bioinformatics.babraham.ac.uk/projects/fastqc

Arabidopsis Genome Initiative (2000). Analysis of the genome sequence of the flowering plant *Arabidopsis thaliana*. *Nature*, *408*(6814), 796–815. https://doi.org/10.1038/35048692

Arnold, J. B., Daroczi, G., Werth, B., Weitzner, B., Kunst, J., Auguie, B., & London, J. (2017). GGTHEMES: *Extra themes, scales and Geoms for "ggplot2" (Version 3.4.0)*. Retrieved from https://cran.r-project.org/web/packages/ggthemes/index.html

Baillie, J., Hilton-Taylor, C., & Stuart, S. N. (2004). *2004 IUCN Red List of threatened species: A global species assessment*. Retrieved from https://portals.iucn.org/library/node/9830

Baquero, O. S. (2017). GGSN: *North symbols and scale bars for maps created with "ggplot2" or "ggmap" (Version 0.4.0)*. Retrieved from https://CRAN.R-project.org/package=ggsn

Benayas, J. M. R., Newton, A. C., Diaz, A., & Bullock, J. M. (2009). Enhancement of biodiversity and ecosystem services by ecological restoration: A meta-analysis. *Science*, *325*(5944), 1121–1124. https://doi.org/10.1126/science.1172460

Bickford, D., Lohman, D. J., Sodhi, N. S., Ng, P. K. L., Meier, R., Winker, K., … Das, I. (2007). Cryptic species as a window on diversity and conservation. *Trends in Ecology & Evolution*, *22*(3), 148–155. https://doi.org/10.1016/j.tree.2006.11.004

Bradbury, I. R., Hamilton, L. C., Dempson, B., Robertson, M. J., Bourret, V., Bernatchez, L., & Verspoor, E. (2015). Transatlantic secondary contact in Atlantic Salmon, comparing microsatellites, a single nucleotide polymorphism array and restriction-site associated DNA sequencing for the resolution of complex spatial structure. *Molecular Ecology*, *24*(20), 5130–5144. https://doi.org/10.1111/mec.13395

Breed, M. F., Stead, M. G., Ottewell, K. M., Gardner, M. G., & Lowe, A. J. (2013). Which provenance and where? Seed sourcing strategies for revegetation in a changing environment. *Conservation Genetics*, *14*(1), 1–10. https://doi.org/10.1007/s10592-012-0425-z

Bucharova, A., Bossdorf, O., Hölzel, N., Kollmann, J., Prasse, R., & Durka, W. (2018). MIX and MATCH: Regional admixture provenancing strikes a balance among different seed-sourcing strategies for ecological restoration. *Conservation Genetics*, *20*(1), 7–17. https://doi.org/10.1007/s10592-018-1067-6

Bucharova, A., Michalski, S., Hermann, J.-M., Heveling, K., Durka, W., Hölzel, N., … Bossdorf, O. (2017). Genetic differentiation and regional adaptation among seed origins used for grassland restoration: Lessons from a multispecies transplant experiment. *Journal of Applied Ecology*, *54*(1), 127–136. https://doi.org/10.1111/1365-2664.12645

Bullock, J. M., Aronson, J., Newton, A. C., Pywell, R. F., & Rey-Benayas, J. M. (2011). Restoration of ecosystem services and biodiversity: Conflicts and opportunities. *Trends in Ecology & Evolution*, *26*(10), 541–549. https://doi.org/10.1016/j.tree.2011.06.011

Burmeier, S., Eckstein, R. L., Donath, T. W., & Otte, A. (2011). Plant pattern development during early post-restoration succession in Grasslands—A case study of Arabis nemorensis. *Restoration Ecology*, *19*(5), 648–659. https://doi.org/10.1111/j.1526-100X.2010.00668.x

Byers, J. E. (2002). Impact of non-indigenous species on natives enhanced by anthropogenic alteration of selection regimes. *Oikos*, *97*(3), 449–458. https://doi.org/10.1034/j.1600-0706.2002.970316.x

Catchen, J., Hohenlohe, P. A., Bassham, S., Amores, A., & Cresko, W. A. (2013). STACKS: An analysis tool set for population genomics. *Molecular Ecology*, *22*(11), 3124–3140. https://doi.org/10.1111/mec.12354

Cavalli-Sforza, L. L., & Edwards, A. W. F. (1967). Phylogenetic analysis. Models and estimation procedures. *American Journal of Human Genetics*, *19*(3 Pt 1), 233–257.

Coiffait-Gombault, C., Buisson, E., & Dutoit, T. (2011). Hay transfer promotes establishment of Mediterranean Steppe vegetation on soil disturbed by pipeline construction. *Restoration Ecology*, *19*(201), 214–222. https://doi.org/10.1111/j.1526-100X.2010.00706.x

Crémieux, L., Bischoff, A., Müller-Schärer, H., & Steinger, T. (2010). Gene flow from foreign provenances into local plant populations: Fitness consequences and implications for biodiversity restoration. *American Journal of Botany*, *97*(1), 94–100. https://doi.org/10.3732/ajb.0900103

Crooks, J. A., Chang, A. L., & Ruiz, G. M. (2011). Aquatic pollution increases the relative success of invasive species. *Biological Invasions*, *13*(1), 165–176. https://doi.org/10.1007/s10530-010-9799-3

Crutsinger, G. M., Collins, M. D., Fordyce, J. A., Gompert, Z., Nice, C. C., & Sanders, N. J. (2006). Plant genotypic diversity predicts community structure and governs an ecosystem process. *Science*, *313*(5789), 966–968. https://doi.org/10.1126/science.1128326

Danecek, P., Auton, A., Abecasis, G., Albers, C. A., Banks, E., DePristo, M. A., … Durbin, R. (2011). The variant call format and VCFtools. *Bioinformatics*, *27*(15), 2156–2158. https://doi.org/10.1093/bioinformatics/btr330

DiLeo, M. F., Rico, Y., Boehmer, H. J., & Wagner, H. H. (2017). An ecological connectivity network maintains genetic diversity of a flagship wildflower, Pulsatilla vulgaris. *Biological Conservation*, *212*, 12–21. https://doi.org/10.1016/j.biocon.2017.05.026

Donath, T. W., Bissels, S., Hölzel, N., & Otte, A. (2007). Large scale application of diaspore transfer with plant material in restoration practice – Impact of seed and microsite limitation. *Biological Conservation*, *138*(1), 224–234. https://doi.org/10.1016/j.biocon.2007.04.020

Edgar, R. C. (2004). MUSCLE: Multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Research*, *32*(5), 1792–1797. https://doi.org/10.1093/nar/gkh340

Elshire, R. J., Glaubitz, J. C., Sun, Q., Poland, J. A., Kawamoto, K., Buckler, E. S., & Mitchell, S. E. (2011). A Robust, Simple Genotyping-by-Sequencing (GBS) approach for high diversity species. *PLoS ONE*, *6*(5), e19379. https://doi.org/10.1371/journal.pone.0019379

Espeland, E. K., Emery, N. C., Mercer, K. L., Woolbright, S. A., Kettenring, K. M., Gepts, P., & Etterson, J. R. (2017). Evolution of plant materials for ecological restoration: Insights from the applied and basic literature. *Journal of Applied Ecology*, *54*(1), 102–115. https://doi.org/10.1111/1365-2664.12739

Etter, P. D., Bassham, S., Hohenlohe, P. A., Johnson, E. A., & Cresko, W. A. (2011). SNP discovery and genotyping for evolutionary genetics using RAD sequencing. *Methods in Molecular Biology*, *772*, 157–178. https://doi.org/10.1007/978-1-61779-228-1_9

Fant, J. B., Holmstrom, R. M., Sirkin, E., Etterson, J. R., & Masi, S. (2008). Genetic structure of threatened native populations and propagules used for restoration in a clonal species, American Beachgrass (*Ammophila breviligulata* Fern.). *Restoration Ecology*, *16*(4), 594–603. https://doi.org/10.1111/j.1526-100X.2007.00348.x

Frankham, R., Ballou, J. D., Eldridge, M. D. B., Lacy, R. C., Ralls, K., Dudash, M. R., & Fenster, C. B. (2011). Predicting the probability of outbreeding depression. *Conservation Biology*, *25*(3), 465–475. https://doi.org/10.1111/j.1523-1739.2011.01662.x

Fukasawa, K., Miyashita, T., Hashimoto, T., Tatara, M., & Abe, S. (2013). Differential population responses of native and alien rodents to an invasive predator, habitat alteration and plant masting. *Proceedings of the Royal Society B: Biological Sciences*, *280*(1773), 20132075. https://doi.org/10.1098/rspb.2013.2075

Gao, S., Bertrand, D., Chia, B. K. H., & Nagarajan, N. (2016). OPERA-LG: Efficient and exact scaffolding of large, repeat-rich eukaryotic genomes with performance guarantees. *Genome Biology*, *17*, 102. https://doi.org/10.1186/s13059-016-0951-y

Gnerre, S., MacCallum, I., Przybylski, D., Ribeiro, F. J., Burton, J. N., Walker, B. J., … Jaffe, D. B. (2011). High-quality draft assemblies of mammalian genomes from massively parallel sequence data. *Proceedings of the National Academy of Sciences of the United States of America*, *108*(4), 1513–1518. https://doi.org/10.1073/pnas.1017351108

Goudet, J., & Jombart, T. (2015). HIERFSTAT: *Estimation and tests of hierarchical F-statistics (Version 0.04-22)*. Retrieved from https://cran.r-project.org/web/packages/hierfstat/index.html

Goulet, B. E., Roda, F., & Hopkins, R. (2017). Hybridization in plants: Old ideas, new techniques. *Plant Physiology*, *173*(1), 65–78. https://doi.org/10.1104/pp.16.01340

Gouy, M., Guindon, S., & Gascuel, O. (2010). SEAVIEW Version 4: A multiplatform graphical user interface for sequence alignment and phylogenetic tree building. *Molecular Biology and Evolution*, *27*(2), 221–224. https://doi.org/10.1093/molbev/msp259

Gruenthal, K. M., Witting, D. A., Ford, T., Neuman, M. J., Williams, J. P., Pondella, D. J., … Larson, W. A. (2014). Development and application of genomic tools to the restoration of green abalone in southern California. *Conservation Genetics*, *15*(1), 109–121. https://doi.org/10.1007/s10592-013-0524-5

Guindon, S., & Gascuel, O. (2003). A simple, fast, and accurate algorithm to estimate large phylogenies by maximum likelihood. *Systematic Biology*, *52*(5), 696–704. https://doi.org/10.1080/10635150390235520

Haas, B. J., Salzberg, S. L., Zhu, W., Pertea, M., Allen, J. E., Orvis, J., … Wortman, J. R. (2008). Automated eukaryotic gene structure annotation using EVidenceModeler and the Program to Assemble Spliced Alignments. *Genome Biology*, *9*(1), R7. https://doi.org/10.1186/gb-2008-9-1-r7

Hand, R., & Gregor, T. (2006). Die Verbreitung von Arabis sagittata in Deutschland. Ergebnisse einer Herbarstudie. *Kochia*, *1*, 21–31.

Haug-Baltzell, A., Stephens, S. A., Davey, S., Scheidegger, C. E., & Lyons, E. (2017). SynMap2 and SynMap3D: Web-based whole-genome synteny browsers. *Bioinformatics*, *33*(14), 2197–2198. https://doi.org/10.1093/bioinformatics/btx144

Hölzel, N., & Otte, A. (2001). The impact of flooding regime on the soil seed bank of flood-meadows. *Journal of Vegetation Science*, *12*(2), 209–218. https://doi.org/10.2307/3236605

Hölzel, N., & Otte, A. (2003). Restoration of a species-rich flood meadow by topsoil removal and diaspore transfer with plant material. *Applied Vegetation Science*, *6*(2), 131–140. https://doi.org/10.1111/j.1654-109X.2003.tb00573.x

Hölzel, N., & Otte, A. (2004). Assessing soil seed bank persistence in flood-meadows: The search for reliable traits. *Journal of Vegetation Science*, *15*(1), 93–100. https://doi.org/10.1111/j.1654-1103.2004.tb02241.x

Hu, T. T., Pattyn, P., Bakker, E. G., Cao, J., Cheng, J.-F., Clark, R. M., … Guo, Y.-L. (2011). The *Arabidopsis lyrata* genome sequence and the basis of rapid genome size change. *Nature Genetics*, *43*(5), 476–481. https://doi.org/10.1038/ng.807

Hufford, K. M., & Mazer, S. J. (2003). Plant ecotypes: Genetic differentiation in the age of ecological restoration. *Trends in Ecology & Evolution*, *18*(3), 147–155. https://doi.org/10.1016/S0169-5347(03)00002-8

Hughes, A. R., Inouye, B. D., Johnson, M. T. J., Underwood, N., & Vellend, M. (2008). Ecological consequences of genetic diversity. *Ecology Letters*, *11*(6), 609–623. https://doi.org/10.1111/j.1461-0248.2008.01179.x

Jalas, J., & Suominen, J. (1994). *Atlas Florae Europaeae. Distribution of Vascular Plants in Europe 10: Cruciferae (Sisymbrium to Aubrieta)*, Vol. *10*. The Committee for Mapping the Flora of Europe & Societas Biologica Fennica Vanamo, Helsinki.

Jeffries, D. L., Copp, G. H., Handley, L. L., Olsén, K. H., Sayer, C. D., & Hänfling, B. (2016). Comparing RADseq and microsatellites to infer complex phylogeographic patterns, an empirical perspective in the Crucian carp, *Carassius carassius* L. *Molecular Ecology*, *25*(13), 2997–3018. https://doi.org/10.1111/mec.13613

Jiao, W.-B., Accinelli, G. G., Hartwig, B., Kiefer, C., Baker, D., Severing, E., … Schneeberger, K. (2017). Improving and correcting the contiguity of long-read genome assemblies of three plant species using optical mapping and chromosome conformation capture data. *Genome Research*, *27*(5), 778–786. https://doi.org/10.1101/gr.213652.116

Jombart, T., Kamvar, Z. N., Lustrik, R., Collins, C., Beugin, M.-P., Knaus, B., & Calboli, F. (2016). ADEGENET: *Exploratory analysis of genetic and genomic data (Version 2.0.1)*. Retrieved from https://cran.r-project.org/web/packages/adegenet/index.html

Kahle, D., & Wickham, H. (2013). GGMAP: Spatial visualization with ggplot2. *The R Journal*, *5*(1), 144–161. https://doi.org/10.32614/RJ-2013-014

Karl, R., & Koch, M. A. (2014). Phylogenetic signatures of adaptation: The *Arabis hirsuta* species aggregate (Brassicaceae) revisited. *Perspectives in Plant Ecology, Evolution and Systematics*, *16*(5), 247–264. https://doi.org/10.1016/j.ppees.2014.06.001

Kiefer, M., Schmickl, R., German, D. A., Mandáková, T., Lysak, M. A., Al-Shehbaz, I. A., … Koch, M. A. (2014). BrassiBase: Introduction to a novel knowledge database on Brassicaceae evolution. *Plant & Cell Physiology*, *55*(1), e3. https://doi.org/10.1093/pcp/pct158

Kiehl, K., Kirmer, A., Donath, T. W., Rasran, L., & Hölzel, N. (2010). Species introduction in restoration projects - evaluation of different techniques for the establishment of semi-natural grasslands in Central

and Northwestern Europe. *Basic and Applied Ecology*, *11*(4), 285–299. https://doi.org/10.1016/j.baae.2009.12.004

Kiehl, K., Kirmer, A., & Shaw, N. (2014). *Guidelines for native seed production and grassland restoration*. Cambridge, UK: Cambridge Scholars Publishing.

Kiehl, K., & Wagner, C. (2006). Effect of Hay transfer on long-term establishment of vegetation and grasshoppers on former arable fields. *Restoration Ecology*, *14*(1), 157–166. https://doi.org/10.1111/j.1526-100X.2006.00116.x

Knaus, B. J., & Grünwald, N. J. (2017). VCFR: A package to manipulate and visualize variant call format data in R. *Molecular Ecology Resources*, *17*(1), 44–53.

Koboldt, D. C., Zhang, Q., Larson, D. E., Shen, D., McLellan, M. D., Lin, L., … Wilson, R. K. (2012). VARSCAN 2: Somatic mutation and copy number alteration discovery in cancer by exome sequencing. *Genome Research*, *22*(3), 568–576. https://doi.org/10.1101/gr.129684.111

Kopelman, N. M., Mayzel, J., Jakobsson, M., Rosenberg, N. A., & Mayrose, I. (2015). CLUMPAK: A program for identifying clustering modes and packaging population structure inferences across K. *Molecular Ecology Resources*, *15*(5), 1179–1191. https://doi.org/10.1111/1755-0998.12387

Korf, I. (2004). Gene finding in novel genomes. *BMC Bioinformatics*, *5*, 59. https://doi.org/10.1186/1471-2105-5-59

Laenen, B., Tedder, A., Nowak, M. D., Toräng, P., Wunder, J., Wötzel, S., … Slotte, T. (2018). Demography and mating system shape the genome-wide impact of purifying selection in Arabis alpina. *Proceedings of the National Academy of Sciences of the United States of America*, *115*, 816–821. https://doi.org/10.1073/pnas.1707492115

Li, H., & Durbin, R. (2009). Fast and accurate short read alignment with Burrows-Wheeler transform. *Bioinformatics*, *25*(14), 1754–1760. https://doi.org/10.1093/bioinformatics/btp324

Li, H., Handsaker, B., Wysoker, A., Fennell, T., Ruan, J., & Homer, N., … 1000 Genome Project Data Processing Subgroup (2009). The sequence Alignment/Map format and SAMtools. *Bioinformatics*, *25*(16), 2078–2079. https://doi.org/10.1093/bioinformatics/btp352

Lyons, E., & Freeling, M. (2008). How to usefully compare homologous plant genes and chromosomes as DNA sequences. *The Plant Journal*, *53*(4), 661–673. https://doi.org/10.1111/j.1365-313X.2007.03326.x

Majoros, W. H., Pertea, M., & Salzberg, S. L. (2004). TIGRSCAN and GLIMMERHMM: Two open source ab initio eukaryotic gene-finders. *Bioinformatics*, *20*(16), 2878–2879. https://doi.org/10.1093/bioinformatics/bth315

Mallet, J., Besansky, N., & Hahn, M. W. (2016). How reticulated are species? *BioEssays*, *38*(2), 140–149. https://doi.org/10.1002/bies.201500149

Martin, M. (2011). Cutadapt removes adapter sequences from high-throughput sequencing reads. *EMBnet.journal*, *17*(1), 10–12. https://doi.org/10.14806/ej.17.1.200

Massatti, R., Doherty, K. D., & Wood, T. E. (2018). Resolving neutral and deterministic contributions to genomic structure in *Syntrichia ruralis* (Bryophyta, Pottiaceae) informs propagule sourcing for dryland restoration. *Conservation Genetics*, *19*(1), 85–97. https://doi.org/10.1007/s10592-017-1026-7

Mattila, T. M., Tyrmi, J., Pyhäjärvi, T., & Savolainen, O. (2017). Genome-wide analysis of colonization history and concomitant selection in *Arabidopsis lyrata*. *Molecular Biology and Evolution*, *34*(10), 2665–2677. https://doi.org/10.1093/molbev/msx193

McKay, J. K., Christian, C. E., Harrison, S., & Rice, K. J. (2005). "How local is local?"—A review of practical and conceptual issues in the genetics of restoration. *Restoration Ecology*, *13*(3), 432–440. https://doi.org/10.1111/j.1526-100X.2005.00058.x

Mijangos, J. L., Pacioni, C., Spencer, P. B. S., & Craig, M. D. (2015). Contribution of genetics to ecological restoration. *Molecular Ecology*, *24*(1), 22–37. https://doi.org/10.1111/mec.12995

Mummenhoff, K., Franzke, A., & Koch, M. (1997). Molecular phylogenetics of Thlaspi s.l. (Brassicaceae) based on chloroplast DNA restriction site variation and sequences of the internal transcribed spacers of nuclear ribosomal DNA. *Canadian Journal of Botany*, *75*(3), 469–482. https://doi.org/10.1139/b97-051

Nei, M. (1987). *Molecular evolutionary genetics*. New York, NY: Columbia University Press.

Novikova, P. Y., Hohmann, N., Nizhynska, V., Tsuchimatsu, T., Ali, J., Muir, G., … Nordborg, M. (2016). Sequencing of the genus *Arabidopsis* identifies a complex history of nonbifurcating speciation and abundant trans-specific polymorphism. *Nature Genetics*, *48*(9), 1077–1082. https://doi.org/10.1038/ng.3617

O'Leary, S. J., Hollenbeck, C. M., Vega, R. R., Gold, J. R., & Portnoy, D. S. (2018). Genetic mapping and comparative genomics to inform restoration enhancement and culture of southern flounder, *Paralichthys lethostigma*. *BMC Genomics*, *19*(1), 163. https://doi.org/10.1186/s12864-018-4541-0

Onge, K. R. S., Källman, T., Slotte, T., Lascoux, M., & Palmé, A. E. (2011). Contrasting demographic history and population structure in *Capsella rubella* and *Capsella grandiflora*, two closely related species with different mating systems. *Molecular Ecology*, *20*(16), 3306–3320. https://doi.org/10.1111/j.1365-294X.2011.05189.x

Paradis, E., Jombart, T., Schliep, K., Potts, A., & Winter, D. (2016). PEGAS: *Population and evolutionary genetics analysis system*. Retrieved from https://cran.r-project.org/web/packages/pegas/index.html

Peterson, B. K., Weber, J. N., Kay, E. H., Fisher, H. S., & Hoekstra, H. E. (2012). Double Digest RADseq: An inexpensive method for de novo SNP discovery and genotyping in model and non-model species. *PLoS ONE*, *7*(5), e37135. https://doi.org/10.1371/journal.pone.0037135

Pfeifer, B., Wittelsbuerger, U., Li, H., & Handsaker, B. (2018). POPGENOME: *An efficient swiss army knife for population genomic analyses (Version 2.6.1)*. Retrieved from https://CRAN.R-project.org/package=PopGenome

Purcell, S. (2009). PLINK *(Version 1.07)*.

Purcell, S., Neale, B., Todd-Brown, K., Thomas, L., Ferreira, M. A. R., Bender, D., … Sham, P. C. (2007). PLINK: A tool set for whole-genome association and population-based linkage analyses. *American Journal of Human Genetics*, *81*(3), 559–575. https://doi.org/10.1086/519795

R Development Core Team (2008). *R: A language and environment for statistical computing*. Vienna, Austria: Foundation for Statistical Computing. Retrieved from http://www.R-project.org

Reed, D. H., & Frankham, R. (2003). Correlation between fitness and genetic diversity. *Conservation Biology*, *17*(1), 230–237. https://doi.org/10.1046/j.1523-1739.2003.01236.x

Reitzel, A. M., Herrera, S., Layden, M. J., Martindale, M. Q., & Shank, T. M. (2013). Going where traditional markers have not gone before: Utility of and promise for RAD sequencing in marine invertebrate phylogeography and population genomics. *Molecular Ecology*, *22*(11), 2953–2970. https://doi.org/10.1111/mec.12228

Reusch, T. B. H., Ehlers, A., Hämmerli, A., & Worm, B. (2005). Ecosystem recovery after climatic extremes enhanced by genotypic diversity. *Proceedings of the National Academy of Sciences of the United States of America*, *102*(8), 2826–2831. https://doi.org/10.1073/pnas.0500008102

Reynolds, L. K., McGlathery, K. J., & Waycott, M. (2012). Genetic diversity enhances restoration success by augmenting ecosystem services. *PLoS ONE*, *7*(6), e38397. https://doi.org/10.1371/journal.pone.0038397

Roberts, L., Stone, R., & Sugden, A. (2009). The rise of restoration ecology. *Science*, *325*(5940), 555. https://doi.org/10.1126/science.325_555

Rochette, N. C., & Catchen, J. M. (2017). Deriving genotypes from RAD-seq short-read data using Stacks. *Nature Protocols*, *12*(12), 2640–2659. https://doi.org/10.1038/nprot.2017.123

Scheepens, J. F., Rauschkolb, R., Ziegler, R., Schroth, V., & Bossdorf, O. (2017). Genotypic diversity and environmental variability affect the invasibility of experimental plant populations. *Oikos*, *127*(4), 570–578. https://doi.org/10.1111/oik.04818

Schnittler, M., & Günther, K.-F. (1999). Central European vascular plants requiring priority conservation measures – An analysis from national Red Lists and distribution maps. *Biodiversity & Conservation*, *8*(7), 891–925. https://doi.org/10.1023/A:1008828704456

Shafer, A. B. A., Peart, C. R., Tusso, S., Maayan, I., Brelsford, A., Wheat, C. W., & Wolf, J. B. W. (2016). Bioinformatic processing of RAD-seq data dramatically impacts downstream population genetic inference. *Methods in Ecology and Evolution*, *8*(8), 907–917. https://doi.org/10.1111/2041-210X.12700

Slater, G. S. C., & Birney, E. (2005). Automated generation of heuristics for biological sequence comparison. *BMC Bioinformatics*, *6*, 31. https://doi.org/10.1186/1471-2105-6-31

Smit, A., & Hubley, R. (2008). *RepeatModeler Open-1.0 (Version 1.0)*. Retrieved from http://www.repeatmasker.org

Smit, A., Hubley, R., & Green, P. (2013). *RepeatMasker Open-4.0*. Retrieved from http://www.repeatmasker.org

Society for Ecological Restoration International Science & Policy Working Group. (2004). *The SER International Primer on Ecological Restoration*. Tucson, AZ: Society for Ecological Restoration International. Retrieved from www.ser.org

Stanke, M., & Waack, S. (2003). Gene prediction with a hidden Markov model and a new intron submodel. *Bioinformatics*, *19*(Suppl 2), ii215–ii225.

Suarez-Gonzalez, A., Lexer, C., & Cronk, Q. C. B. (2018). Adaptive introgression: A plant perspective. *Biology Letters*, *14*(3), 20170688. https://doi.org/10.1098/rsbl.2017.0688

Suding, K. N. (2011). Toward an era of restoration in ecology: Successes, failures, and opportunities ahead. *Annual Review of Ecology, Evolution, and Systematics*, *42*(1), 465–487. https://doi.org/10.1146/annurev-ecolsys-102710-145115

Tellier, A., Laurent, S. J. Y., Lainer, H., Pavlidis, P., & Stephan, W. (2011). Inference of seed bank parameters in two wild tomato species using ecological and genetic data. *Proceedings of the National Academy of Sciences of the United States of America*, *108*(41), 17052–17057. https://doi.org/10.1073/pnas.1111266108

Török, P., Miglécz, T., Valkó, O., Kelemen, A., Tóth, K., Lengyel, S., & Tóthmérész, B. (2012). Fast restoration of grassland vegetation by a combination of seed mixture sowing and low-diversity hay transfer. *Ecological Engineering*, *44*, 133–138. https://doi.org/10.1016/j.ecoleng.2012.03.010

Torres-Martinez, L., & Emery, N. C. (2016). Genome-wide SNP discovery in the annual herb, Lasthenia fremontii (Asteraceae): Genetic resources for the conservation and restoration of a California vernal pool endemic. *Conservation Genetics Resources*, *8*(2), 145–158. https://doi.org/10.1007/s12686-016-0524-0

Tyrrell, M. C., & Byers, J. E. (2007). Do artificial substrates favor nonindigenous fouling species over native species? *Journal of Experimental Marine Biology and Ecology*, *342*(1), 54–60. https://doi.org/10.1016/j.jembe.2006.10.014

Vandepitte, K., Gristina, A. S., Hert, K. D., Meekers, T., Roldán-Ruiz, I., & Honnay, O. (2012). Recolonization after habitat restoration leads to decreased genetic variation in populations of a terrestrial orchid. *Molecular Ecology*, *21*(17), 4206–4215. https://doi.org/10.1111/j.1365-294X.2012.05698.x

Vander Mijnsbrugge, K., Bischoff, A., & Smith, B. (2010). A question of origin: Where and how to collect seed for ecological restoration. *Basic and Applied Ecology*, *11*(4), 300–311. https://doi.org/10.1016/j.baae.2009.09.002

Vrijenhoek, R. C. (1994). Genetic diversity and fitness in small populations. In V. Loeschcke, S. K. Jain, & J. Tomiuk (Eds.), *Conservation genetics* (pp. 37–53). Basel, Switzerland: Birkhäuser.

Weeks, A. R., Sgro, C. M., Young, A. G., Frankham, R., Mitchell, N. J., Miller, K. A., … Hoffmann, A. A. (2011). Assessing the benefits and risks of translocations in changing environments: A genetic perspective. *Evolutionary Applications*, *4*(6), 709–725. https://doi.org/10.1111/j.1752-4571.2011.00192.x

Wickham, H. (2009). GGPLOT*2: Elegant graphics for data analysis*. Retrieved from http://ggplot2.org

Williams, A. V., Nevill, P. G., & Krauss, S. L. (2014). Next generation restoration genetics: Applications and opportunities. *Trends in Plant Science*, *19*(8), 529–537. https://doi.org/10.1016/j.tplants.2014.03.011

Willing, E.-M., Rawat, V., Mandáková, T., Maumus, F., James, G. V., Nordström, K. J. V., … Schneeberger, K. (2015). Genome expansion of Arabis alpina linked with retrotransposition and reduced symmetric DNA methylation. *Nature Plants*, *1*(2), 14023. https://doi.org/10.1038/nplants.2014.23

Zhao, S., Guo, Y., Sheng, Q., & Shyr, Y. (2015). HEATMAP*3: An Improved Heatmap Package (Version 1.1.1)*. Retrieved from https://cran.r-project.org/web/packages/heatmap3/index.html

SUPPORTING INFORMATION

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

How to cite this article: Dittberner H, Becker C, Jiao W-B, et al. Strengths and potential pitfalls of hay transfer for ecological restoration revealed by RAD-seq analysis in floodplain *Arabis* species. *Mol Ecol*. 2019;28:3887–3901. https://doi.org/10.1111/mec.15194