Intact rDNA arrays of *Potentilla*-origin detected in *Erythronium* nucleus suggest recent eudicot-to-monocot horizontal transfer

László Bartha[a][1], Terezie Mandáková[b], Aleš Kovařík[c], Paul-Adrian Bulzu[d], Nathalie Rodde[e], Václav Mahelka[f], Martin A. Lysak[b], Margaux-Alison Fustier[e], Jan Šafář[g], Petr Cápal[g], Lujza Keresztes[h,i], Horia L. Banciu[a,i,j]

[a]Molecular Biology Center, Interdisciplinary Research Institute on Bio-Nano-Sciences, Babeș-Bolyai University, 400271 Cluj-Napoca, Romania

[b]Central European Institute of Technology (CEITEC), Masaryk University, 625 00 Brno, Czech Republic

[c]Laboratory of Molecular Epigenetics, Institute of Biophysics, Academy of Sciences of the Czech Republic, 61265 Brno, Czech Republic

[d]Department of Aquatic Microbial Ecology, Institute of Hydrobiology, Biology Centre of the Academy of Sciences of the Czech Republic, 370 05 České Budějovice, Czech Republic

[e]French Plant Genomic Resource Center, INRAE-CNRGV, 31320 Castanet Tolosan, France

[f]Institute of Botany, Czech Academy of Sciences, 25243 Průhonice, Czech Republic

[g]Institute of Experimental Botany of the Czech Academy of Sciences, Centre of the Region Haná for Biotechnological and Agricultural Research, 779 00 Olomouc, Czech Republic

[h]Hungarian Department of Biology and Ecology, Babeș-Bolyai University, 400006 Cluj-Napoca, Romania

[i]Centre of Systems Biology, Biodiversity and Bioresources (3B), Babeș-Bolyai University, 400006 Cluj-Napoca, Romania

[j]Department of Molecular Biology and Biotechnology, Faculty of Biology and Geology, Babeş-Bolyai University, 400006 Cluj-Napoca, Romania

[1]To whom correspondence may be addressed. **Email: *lbartha.ubbcluj@yahoo.com***

1

Classification

Biological Sciences. Plant Biology.

Keywords

eudicot, monocot, rDNA, ribosomal RNA genes, lateral gene transfer.

Author Contributions

L.B., T.M., A.K., N.R., V.M., M.A.L., J.Š., L.K., and H.L.B. designed research; L.B., T.M., A.K., P.A.B., N.R., J.Š., and P.C. performed research; A.K., N.R., M.A.L., J.Š., L.K., and H.L.B. contributed new reagents/analytic tools; L.B., T.M., A.K., P.A.B., M.A.L., and M.A.F. analyzed data; L.B. and V.M. drafted the manuscript; all co-authors contributed to the final version of manuscript.

Data deposition: The sequences reported in this paper have been deposited in the GenBank database [accession nos. MW357750, MW357751 (BAC clones), MW400601–MW400606, MW357749 (ITS/ETS sequences), and MW930739–MW930743 (ITS2 sequences)].

Competing Interest Statement: The authors declare no conflict of interest.

This PDF file includes:

Main Text
Figures 1 to 5

Abstract.

The occurrence of horizontal gene transfer (HGT) in Eukarya is increasingly gaining recognition. Nuclear-to-nuclear jump of DNA between plant species at high phylogenetic distance and devoid of intimate association (e.g., parasitism) is still scarcely reported. Within eukaryotes, components of ribosomal DNA (rDNA) multigene family have been found to be horizontally transferred in protists, fungi and grasses. However, in neither case HGT occurred between phylogenetic families, nor the transferred rDNA remained tandemly arrayed and transcriptionally active in the recipient organism. This study aimed to characterize an alien eudicot-type of 45S nuclear rDNA, assumingly transferred horizontally to the genome of monocot European *Erythronium* (Liliaceae). Genome skimming coupled by PacBio HiFi sequencing of a BAC clone were applied to determine DNA sequence of the alien rDNA. A clear phylogenetic signal traced the origin of the alien rDNA of *Erythronium* back to the Argentea clade of *Potentilla* (Rosaceae) and deemed the transfer to have occurred in the common ancestor of *E. dens-canis* and *E. caucasicum*. Though being discontinuous, transferred rDNA preserved its general tandemly arrayed feature in the host organism. Southern blotting, molecular cytogenetics, and sequencing of a BAC clone derived from flow-sorted nuclei indicated integration of the alien rDNA into the recipient's nuclear genome. Unprecedently, dicot-type alien rDNA was found to be transcribed in the monocot *Erythronium* albeit much less efficiently than the native counterpart. This study adds a new example to the growing list of naturally transgenic plants while holding the scientific community continually in suspense about the mode of DNA transfer.

Significance Statement.

Ribosomal DNA is an essential component of all cellular genomes. In plants, accidental movement of rDNA via horizontal gene transfer has only been reported in sexually incompatible grasses (monocots) where it involved non-functional rDNA units. In this study, we propose that evolutionary trajectories of eudicots and monocots were bypassed by the jump of rDNA from a *Potentilla* species (Rosaceae) to a common ancestor of *Erythronium dens-canis* and *E. caucasicum* (Liliaceae). The alien eudicot-type rDNA appeared relatively well conserved in the examined host *Erythronium* genome, being able to be expressed while preserving its general tandemly repeated feature, evidences that have no match in earlier literature.

Main Text

Introduction

Horizontal gene transfer (or alternatively, lateral gene transfer, LGT) is an evolutionary mechanism surmounting reproductive barriers between species. It refers to the movement of genetic material between organisms through routes other than parent-to-offspring. The phenomenon is widespread in prokaryotes and its significance and prevalence are also increasingly recognized in eukaryotes (1, 2). Successful gene transfers can result in evolutionary innovations, like adaptation to a new environment (3), improvement of parasitic lifestyle (4), or a shift in metabolism resulting in C4 biochemistry (5).

Concerning plant-to-plant gene transfers, the very first examples of HGT referred mostly to the jump of mitochondrial DNAs within host-parasite systems [see review (6)]. Later, reporting HGT of nuclear genes and their expression in parasitic plants also became mainstream (7). Not surprisingly, the host-parasite interface provides a suitable opportunity for gene movement via the intimate physical connections between donor and recipient organisms. More recently, HGT was shown to frequently occur in certain groups of plants like grasses where it was evidenced that phylogenetic proximity and rhizomatous nature increased the frequency of HGT (8).

Nucleus-to-nucleus jump of DNA among angiosperms that are found at high phylogenetic distance (e.g., those belonging to different families) and not showing some kind of biological promiscuity (e.g., parasitism) are still rarely reported. Indeed, transfer of a nuclear-encoded transposable element (TE) between the dicot grapevine and the monocot palm tree has been reported (9). These genomic components are thought to be particularly prone to HGT thanks to their propensity for self-replication, extrachromosomal stage from within their transposition cycle, and the capacity to integrate themselves into host chromosomes (10-12).

Ribosomal DNA constitutes an essential component of all cellular genomes. It encodes ribosomal RNA (rRNA) molecules that together with ribosomal proteins make up ribosomes, factories for protein synthesis. In the majority of angiosperms, three of the RNA genes [18S, 5.8S, and 25S (26S)] are physically linked in a conserved cluster referred to as 45S rDNA. The above genes are separated by two internal transcribed spacers (ITS1 and ITS2) and altogether form a single transcription unit (13, 14). The units are arrayed tandemly and are separated by the so-called intergenic spacer (IGS) region, the 3'-end of which is commonly referred to as external transcribed spacer (ETS) (14).

In plants, intra-individual heterogeneity of rDNA is a well-documented feature. Withing a genome, different rDNA units (e.g., those resulting from hybridization or pseudogenization) can be homogenized by concerted evolution via unequal crossing-over and gene conversion (15, 16). The propensity to maintain or erase intra-individual rDNA heterogeneity/paralogy seems to vary across genomic and taxonomic contexts. As such, rDNA can be particularly polymorph in ancient plant lineages (17, 18) but in the same time, it can be subjected to rapid unidirectional gene conversion in recently formed allopolyploids (19, 20).

According to the 'complexity hypothesis' (21), members of complexes involved in the transcriptional and translational processes (a.k.a. 'informational genes') are less prone to HGT when compared to the 'operational genes' (i.e. 'metabolic housekeeping genes'). Indeed, albeit HGT of rDNA has been repeatedly documented in prokaryotes (22-24), it was shown to occur at a low rate and generally was limited to take place between closely related taxa upon a survey of 2143 prokaryotic genomes (24). Within eukaryotes, HGT of 45S rDNA elements has been revealed among protists (25), fungi (26) and monocot plants (grasses) (27). Published evidences of horizontally transferred eukaryotic rDNA concern only jumps at the intrafamily level and, if tested, in none case the transferred rDNA was shown to be transcriptionally active in the recipient organism. Moreover, the general tandemly arrayed nature of eukaryotic 45S rDNA (28) was not yet confirmed in any host organism after its transfer.

The genus *Erythronium* belongs to the monocot family Liliaceae and includes ca. 30 species of bulbous geophytes confined mainly to temperate forest and damp meadow habitats of the northern hemisphere (29, 30). Species of the genus cluster in three well supported and biogeographically meaningful (eastern North American, western North American and Eurasian) lineages (29, 30). During a phylogenetic study (31) of Eurasian *Erythronium*, Sanger sequencing of the ITS region of 45S nuclear rDNA in *Erythronium dens-canis* L. did not provide

sequences of satisfactory quality. This led to the necessity of cloning of the ITS amplicons. Surprisingly, upon BLAST searches, a subset of the sequenced clones showed 97 percent identity with ITS sequences of the eudicot (hereafter for simplicity named dicot) genus *Potentilla* (Rosaceae). This finding hinted at either a possible artefact, or the actual presence of *Potentilla*-type DNA within the genome of *Erythronium*. A series of experiments conducted subsequently in different laboratories supported the latter scenario, providing evidence for sharing genetic material between plant lineages which separated ca. 152 Mya (http://www.timetree.org/) (32). Hence, in this study, we summarize the results of analyses aiming to characterize the nature of *Potentilla*-specific rDNA within the genome of the monocot plant *Erythronium*. Particular challenges of documenting alien rDNA in the host genome were raised by the general repetitive nature of rDNA and by the large genome of *Erythronium* combined with the relative low amount of alien rDNA therein. A previous genome size measurement of the species estimated a monoploid size (1Cx) to approx. 12.2 gigabase pairs (Gbp) (33), which was more than twice the mean angiosperm 1C value reported recently (34).

Results and Discussion

rDNA Jumped Between Plant Families.

Throughout the study, for consistency purposes an effort was made to conduct key analyses on specimens from a single population of *E. dens-canis*. For this purpose, we selected the so-called 'Feleacu' population from NW Romania (Table S1). Endeavors to differentially PCR-amplify *Potentilla*-specific ITS2 in selected New and Old World *Erythronium* taxa resulted in the expected product only in case of *E. dens-canis* and *E. caucasicum* Woronow (Fig. S1). Besides the initial cloning of ITS in *E. dens-canis*, additional strategies for retrieving and tracing back the origin of the alien rDNA included *de novo* and reference-guided assemblies of ITS and partial ETS from genomic Illumina reads. Furthermore, a low-coverage bacterial artificial chromosome (BAC) library was constructed from flow-sorted nuclei of *E. dens-canis* and screened for *Potentilla*- and *Erythronium*-specific ITS2. The two sequenced BAC clones, 1J24 and 18I01, contained alien and native rDNA, respectively, from which ITS and partial ETS sequences were extracted for phylogeny purposes. Alien-type ITS sequences of *Erythronium* showed 97% identity with *Potentilla* in BLAST search, pointing to this genus as a likely source of the sequences. A follow-up ML phylogenetic analysis placed the sequences with maximal statistical support within the Argentea clade of the genus (Fig. 1).

The native *E. dens-canis* ITS from BAC clone 18I01 fitted well the previously published Eurasian *Erythronium* phylogeny (31), serving as a general control (Fig. S2). In nuclear ITS and plastid DNA-based phylogenies of the genus, *E. dens-canis* and *E. caucasicum* were consistently resolved as sister taxa (29-31). Additionally, populations of *E. dens-canis* subdivide into the so-called 'Transylvanian' and 'non-Transylvanian' phylogeographic lineages based on plastid DNA (35). The 'Feleacu' population belongs to the 'Transylvanian lineage' of the species. In order to enlarge the coverage area of the study, four *E. dens-canis* specimens from different western and central European populations [all belonging to the 'non-Transylvanian lineage' (35)] were additionally spot-checked for *Potentilla*-type ITS2. In all cases the presence of alien ITS2 was confirmed by direct sequencing (Table S1). The phylogenetically meaningful presence of *Potentilla*-specific rDNA in European *Erythronium* can be best interpreted as a HGT event preceding the split of *E. dens-canis* and *E. caucasicum*. One can speculate that alien rDNA is present in most of the populations of *E. dens-canis* across its latitudinally narrow and longitudinally wide European range. Testing this assumption, however, was beyond the scope of the present study.

The invoked ML phylogeny confidently traced the origin of alien rDNA back to the donor organism at the generic level. *Potentilla* is a species-rich genus comprising ~400 typically yellow-flowered representatives in the Northern Hemisphere (36). Most of the members (~300 spp.) belong to the informally named Argentea clade (37), in which species relationships are generally unresolved [(38), Fig. 1]. Illumina consensus sequences of alien ITS from *E. dens-canis* and *E. caucasicum* are identical, pointing to their recent acquisition from *Potentilla* and conserved nature in *Erythronium*. Since the sequences do not cluster with statistical significance with any current *Potentilla* species included into the analysis, the donor remains unknown. Revealing the donor's identity or narrowing relationships of the donor organism are

5

challenging tasks, that can potentially be addressed in the future with the aid of more variable genetic markers and/or by increasing sampling of ITS sequences from the genus of *Potentilla*.

Alien rDNA is integrated in the nuclear genome of *Erythronium*.
Importantly, a series of experiments was conducted to unambiguously prove the insertion of alien rDNA in the host *Erythronium* genome. The experiments led to the elimination of both possible contamination of DNA or samples, and potential involvement of organellar genomes as the locations of alien DNA.

A Southern blot technique was employed using *Apo*I-digested genomic DNAs of samples of *E. dens-canis* as the host genome, complemented by *Erythronium sibiricum* subsp. *altaicum* Rukšāns (later for simplicity mentioned as *E. sibiricum*) used as a negative control and *Potentilla reptans* L. used as a positive control. Digested DNA of *E. dens-canis* produced a distinct band of ~0.9 kilobase pairs (kbp) when hybridized to either *Potentilla*-specific ITS1-5.8S-ITS2 or ITS2 probes. This result fitted expectations based on the restriction map (Fig. 2). This band of *Potentilla*-origin was missing in *E. sibiricum*. At the applied default stringency recommended by the labeling kit producer, probes were not able to discriminate between the native and alien rDNA fragments resulting from the digestion. Specificity was therefore guaranteed by differences in band sizes. Due to the high sequence conservation of 5.8S regions between *Erythronium* and *Potentilla*, binding of the longer probe also to the ~2.4 kbp long native rDNA fragment of *Erythronium* was expected. On the other hand, hybridization of *Potentilla* ITS2 probe to this fragment was somehow surprising. It seems that the 51% sequence identity (over 260 bp) between *Erythronium* and *Potentilla* ITS2 hindered the design of a fully specific probe. At least the bands are more clear-cut in the case of the shorter probe, indicating an improvement in specificity. Smears in case of *E. sibiricum* may indicate that its rDNA is not homogenous and does not conform entirely to the predicted digestion pattern.

Additional evidence of nuclear insertion of alien rDNA was provided by the alien rDNA-containing BAC clone 1J24. BAC library of *E. dens-canis* was constructed from flow-sorted nuclei, eliminating the possibility of *Potentilla*-rDNA to represent any kind of artefact.

Our efforts to directly visualize parts of alien rDNA using fluorescence *in situ* hybridization (FISH) technique failed. Probes either cross-hybridized with native loci (in case of ITS1 and ETS) or provided no signals (in case of 5'-end IGS). We therefore resorted to an indirect approach targeting native ETS and, additionally, native-alien 26S gene on *E. dens-canis* and *E. sibiricum* chromosomes. The 26S probe identified two minor terminal loci on two different, non-homologous chromosomes of *E. dens-canis* that were extraneous to the native ETS loci (Fig. 3). These can be interpreted as alien rDNA. Such additional loci were not observed in *E. sibiricum* (Fig. S3). We are aware of the fact that chromosomal locations in which native and alien rDNA co-localize could not be identified by this approach therefore the real number of alien rDNA loci could be higher.

Alien Dicot rDNA Is Organized Tandemly in the Host Monocot Genome.
Physical map of BAC clone 1J24 provides unprecedented insight into the organization of a dicot rDNA within the genome of a monocot plant (Fig. 4). BAC clone 1J24 contained eight transcription units of *Potentilla*-derived rDNA. Five out of the eight copies of 26S rRNA gene are disrupted by potentially coding sequences (CDSs) of opposite orientation. In one additional case, the 26S rDNA gene is interrupted by an insertion of sequence of unknown origin. The CDSs are not identical and all contain the domain of unknown function (DUF) 4283. Interestingly, similar hypothetical CDSs interleaving 26S rRNA gene were also found within a considerably longer array of the native rDNA of *E. dens-canis* found within the 125.9 kbp long BAC clone 18I01 (Fig. S4).

A Neighbor-Joining tree was generated based on the alignment of translated DUF4283 sequences of the two BAC clones supplemented with publicly available similar protein sequences (Fig. S5). Sequences belonging to BAC 1J24 formed a clade to which the domain sequence of BAC 18I01 was resolved as sister. This suggests their homology. Unfortunately, the phylogeny is not robust enough to establish the phylogenetic origin (dicot versus monocot) of the BAC CDSs. Most of the participating sequences are hypothetical and/or uncharacterized, leaving us with a conundrum about the possible roles of these CDSs in the *Erythronium* genome.

6

In addition to potential CDSs, short (up to 531 bp long) transposon-like fragments were also identified in BAC clone 1J24 (Table S2). These, however, did not assemble (at least in this clone) into complete transposable elements. TEs have been the subject of HGT in plants relatively frequently and are viewed as notably prone to HGT (9-12, 39). Whether this relies on the fact that they exist as extrachromosomal elements during their transposition cycle, remains to be clarified (1). TEs are thought to be able to insert themselves into vector DNA and, subsequently, into host chromosomes (10). Once integrated into the host genome, they can undergo various processes, like mutation, epigenetic silencing, relocation or amplification (transpositional burst) (12). As such, they can reshuffle genomic regions (40). The fact that TEs often accompany foreign DNA in the host genomes (27, 41, 42) suggests that they could play a role in the DNA transfer. Moreover, DNA fragments similar to the *Potentilla*-sister *Fragaria* DNA transposon CACTA elements found within BAC clone 1J24 (Table S2) suggest transposable elements might have also been participating in the actual transfer.

Simultaneous presence of native and alien rDNA was not detected in the sequenced BAC clones although a chimeric sequence read containing parts of both native and alien rDNA (Fig. S6 and Dataset S1) was found within the Nanopore whole genome shotgun (WGS) dataset (Tables S3, S4). The fragmentation of native rDNA observed in this sequence is possibly result of pseudogenization. Nevertheless, given the rarity of this sequence type, one cannot rule out its origin as a sequencing artefact. Presumed pseudogenization of alien rDNA at a larger scale was caught in action in the longest Nanopore read containing *Potentilla*-specific rDNA (Fig. S7 and Dataset S2). Here, disruption of rDNA subunits has a more or less symmetrical pattern within a palindromic structure that renders little likelihood for the artefact origin. Numbers of polymorphic nucleotide sites were also assessed in alignments of native and alien 18S and 26S sequences from within the two analyzed BAC clones. Pseudogenization of alien rDNA is further supported by the fact that number of polymorphic nucleotide sites is several folds higher in alien 18S and 26S sequences when compared to their native homologues (Table S5). Still, conclusions in this respect must be taken cautiously since nucleotide variation in two BAC clones cannot truly reflect variation of target regions across the whole-genome level.

Albeit the presence of horizontally-transferred rDNA was demonstrated previously in eukaryotes (25, 27, 43), in none of the cases the detected alien DNA was arrayed like shown in the present study.

Alien rDNA Is Expressed in *Erythronium*.
The presence of tandemly arrayed alien rDNA units in the *Erythronium* genome raised the possibility of their transcription. To test this supposition, RNA was extracted from actively growing roots of a cultivated *E. dens-canis* plant from the 'Feleacu' population, reverse-transcribed, and subjected to a regular PCR by using *Potentilla* ITS2-specific primers (Table S6). The experiment showed that the alien rDNA was transcriptionally active in *Erythronium* roots (Fig. 5). This finding, however, does not necessarily mean that expressed *Potentilla*-specific rRNA is 'biologically active' in the sense of taking part in the protein synthesis machinery of the host plant. Alien and native 18S and 26S RNAs might be too diverged (95.7 and 90.5 percent identities, respectively) to be interchangeable within a fully functional ribosome. Ribosomal heterogeneity is a well-documented concept and although it can be defined by rRNA sequence variation, the exhaustive list of previously published examples covers mainly heterogeneity determined by ribosomal protein paralogy (13). No precedent in the literature can be found about 18S or 26S rRNA variations leading to deviation of canonical ribosomal structure or differences in translation efficiencies between the uses of different rRNAs by ribosomes of the same plant. In any case, transcription of assumedly horizontally transferred rDNA found within the current study is unprecedented in the literature.

Another issue that was intriguing to explore was the amount of the alien rDNA in the host genome. Specifically, we measured relative fold difference (RFD) between the native and its alien homologue of rDNA within the host genome and transcriptome. A semi-quantitative PCR (SQ-PCR) approach was carried out to perform pilot estimates of this ratio in *E. caucasicum* gDNA, as well as in *E. dens-canis* gDNA and cDNA. Within gDNA, highest values of RFD (553) were found in *E. caucasicum* and in one accession of *E. dens-canis* from the eastern part of the species' range (i.e., the 'Feleacu' population) (Tables S1, S7). RFD tended to decrease towards the western end of species' range, reaching values of 195 and 88 in the analyzed Spanish and French samples, respectively, while being somewhat intermediate (259) in the

7

Croatian sample. Ideally, a future population phylogenomic study should establish whether *E. dens-canis* had an east-west spread across Southern Europe. In such a case, estimates of RFD accompanied by genome size measurements on a larger number of populations should be performed to test whether RFD between the monocot and dicot rDNA changed in the genome of *E. dens-canis* across a longitudinal geographic gradient. Such study may provide novel insights into the dynamics of evolution of rDNA in plant populations across space and time.

Within cDNA of one sample of the core *E. dens-canis* population analyzed, there appeared to be approximately 3000 times less alien ITS2 when compared to its native counterpart (Table S7). This finding – in combination with quantity estimates based on gDNA – signals that transcription efficiency of dicot rDNA in *Erythronium* is several folds lower than that of the native loci. Predicting factors that contributed to this transcription inefficiency (beyond the already mentioned pseudogenization) would be mere speculation. The ongoing Iso-Seq analysis of *E. dens-canis* RNA will potentially shed further light on the transcription efficiency of presumably horizontally transferred rDNA in plants.

Mode of DNA Transfer Remains Elusive.
An intriguing question that generally arises in the context of HGT concerns the mode of the transfer. In this study, total length of the alien DNA of *Potentilla* origin is not known. Nevertheless, the presence of stretches of alien rDNA spanning at least 85.4 kbp in case of a BAC clone and 67.8 kbp in case of Nanopore reads suggests that alien DNA may amount to hundreds of kilobases. Transfers of such pieces of DNA are not exceptional (27, 41, 44). Given the potential large size of transferred DNA, one tends to favor a direct transfer via, for example, a sap-sucking insect over a delivery via a vector organism with relatively small genome size (like bacteria or virus) into whose genome incorporation of alien DNA should have occurred first. Due to this reason a direct transfer seems to be more parsimonious. Parasitic association was shown to increase the chance of HGT (4, 39, 44-47). To our best knowledge, no insect parasite shared by *Erythronium* and *Potentilla* is known to date. It is well documented that *Erythronium* is parasitized by the rust fungus *Uromyces erythronii* (DC.) Pass. in Eurasia (48, 49). A survey of European species belonging to Rosaceae identified 39 *Potentilla* species as hosts of rust fungi (50). A surprisingly large haploid genome size of 2489 Mbp has been reported in another *Uromyces* species (51) and genome sizes are generally high in rust fungi (Pucciniales) (52). Taken together, transfer of alien DNA into the *Erythronium* genome by a rust fungus cannot be excluded though it contradicts the direct transfer hypothesis favored above.

Erythronium is a textbook example of myrmecochory (dispersal of elaiosome-bearing seeds by ants) and *Potentilla* is one of the two genera of Rosaceae identified so far in which myrmecochory also was reported (53). If myrmecochory somehow contributed to the HGT event, integration of alien DNA into the *Erythronium* genome could not occur prior to seed ripening. Recently, within grasses, a statistical increase of the number of HGTs was found in rhizomatous species (8). Perennial plants able to regenerate from vegetative parts (like rhizomes or bulbs) may have an increased opportunity to fix alien DNA in their germline which was previously transferred e.g., via cell-cell contact (8). *Erythronium* species are perennials and vegetative propagation of bulbs is common in the genus. Therefore, alien DNA potentially captured during e.g., germination stage could be transferred to a whole new plant via bulb formation. Such scenario also fits the weak-link model of HGT (54).

Along with direct transfers based on cell-cell contact, genetic material exchange following non-standard (illegitimate) pollination has been postulated in grasses (55). This scenario presumes growth of *Potentilla* pollen tube on *Erythronium* stigma to enable exchange of DNA between chromosomes. Such process would require certain degree of compatibility. So far, regular wide hybridization involving *Potentilla* pollen was documented with *Fragaria* (Rosaceae), leading to their commonly reported intergeneric hybridization (56). Due to the phylogenetic distance of the monocot and dicot lineages, the above scenario of HGT via pollination seems to be unlikely. The current study adds a new example to the growing list of naturally transgenic plants and essentially suggests that 'in plants anything is possible'.

Conclusions and Perspectives.
In this article, authors invoked the horizontal gene transfer phenomenon to explain the presence of tandemly arrayed and transcriptionally active dicot-type rDNA in the nuclear genome of a

monocot plant. Working with different parts (leaves, bulbs, root tips) of the host *Erythronium* plant in different laboratories refuted an artefact origin of the alien rDNA. *Potentilla*-type ITS and ETS sequences perfectly fitting a clear phylogenetic framework of the genus *Potentilla*, coupled by >150 My separate evolution of dicot and monocot lineages represent indirect but strong arguments in support of HGT scenario.

The underlying evolutionary accident between a dicot and monocot plant might had consequences for the *Erythronium* genome reaching far beyond the simple hosting of an alien-type rDNA. *Erythronium* genome will have to be carefully tiptoed in search for non-rDNA sequences of *Potentilla* origin. During this, particular attention should be paid to transposable elements and their role in the transfer, either as passive hitchhikers, or active agents, carriers, or rebuilders. Should any coding sequences be discovered, investigating their processing as well as functionality at the background of monocot genomic apparatus would be unprecedented. Ramping availability of genomic resources will provide ample opportunity to determine the extent of HGT with all of its consequences to the *Erythronium* genome.

Materials and Methods

Plant Materials and DNA Extraction. The list of *Erythronium* samples analyzed in the study (including geographical origin, voucher information and specific molecular analyses used) is shown in Table S1. DNA was preferentially extracted from fresh bulbs using DNeasy Plant Mini Kit (Qiagen, Hilden, Germany). Pilot molecular cytogenetic analyses confirmed specimens of the main analyzed *E. dens-canis* ('Feleacu') population to be diploid (2n=24), according to the common ploidy level reported for the species (57). For certain techniques where inclusion of a positive control was necessary, fresh leaves of *Potentilla reptans* L. were processed for DNA extraction.

PCR Amplifications and Sanger Sequencing. MyTaq™ Red Mix (Bioline, London, UK) was utilized for PCR amplifications throughout the study. Primers ITS5 and ITS4 (58) were used for routine amplification of rDNA ITS. For differential amplification of *Potentilla*-specific ITS2 in *Erythronium*, new primers were designed (Table S6). The single instance of cloning of *E. dens-canis* ITS region was performed as described previously (31). *Potentilla reptans* ITS, *Erythronium* ITS clones and *Potentilla* ITS2-specific PCR products of *Erythronium* were Sanger sequenced at Macrogen Europe B.V. (Amsterdam, The Netherlands).

Phylogenetic Analyses. *Potentilla*-type ITS and partial ETS sequences retrieved from BAC clones and Illumina datasets of *Erythronium* were added into a recent backbone phylogeny of *Potentilla* (59). This phylogeny was supplemented with additional *Potentilla* taxa having publicly available ITS sequences that were potentially closely related to the alien ITS of *Erythronium*. A maximum likelihood (ML) phylogeny was computed with the IQ-TREE (60) web server (http://iqtree.cibiv.univie.ac.at/). The strategy for public *Potentilla* ITS sequence gathering as well as details of their aligning and parameters of the applied ML phylogeny are provided in *SI Materials and Methods*. GenBank accession numbers for newly generated and previously published *Potentilla* and *Erythronium* sequences included into the study are listed in Tables S1, S8 and S9, respectively.

Southern Blot Analysis. In addition to constructing a BAC library from flow-sorted nuclei (see further), Southern blotting technique was used to verify integration of dicot rDNA into the genome of *E. dens-canis*. Based on the BAC clone sequences of *E. dens-canis*, it was predicted the *Apo*I restriction enzyme to produce an ~878 bp long fragment selectively from the dicot-type rDNA. This comprised 3'-end of 18S rRNA gene, ITS1, 5.8S rRNA gene and ITS2. In contrast, a 2468 bp long fragment was expected to arise from the native 45S rDNA of *Erythronium* that also contained complete ITS region (Fig. 2a). Forty-five micrograms of *Erythronium* and five micrograms of *Potentilla* gDNAs (concentrated by ethanol precipitation to 0.5 µg/µl) were digested with *Apo*I-HF enzyme (New England Biolabs, Hertfordshire, UK), loaded onto 1% agarose gel, and electrophoretically separated overnight. Size-selected DNA fragments were transferred onto Amersham Hybond-N+ nylon membrane (GE Healthcare, Buckinghamshire, UK) by high salt transfer with 20×SSC, followed by UV-crosslinking (0.120 J/cm^2). Amersham Alkphos Direct labeling reagents (GE Healthcare, Buckinghamshire, UK) were used for probe labeling and subsequent hybridization. The blot was probed twice. At first, gel-purified ITS region of *P. reptans* amplified from gDNA was used as probe. For the second time, *Potentilla*-specific ITS2 amplified from a dicot-type ITS clone of *E. dens-canis* which served as probe. Conditions for hybridization and post hybridization washes were the same as in Amersham protocol. Hybridization lasted overnight in a 30 cm long hybridization tube containing 30 ml hybridization buffer and 300 ng of probe DNA. Chemiluminescence detection was achieved with CDP-*Star* in a ChemiDoc MP imaging system (BioRad, Hercules, CA, USA).

Chromosome Preparation. Mitotic chromosome spreads were prepared from root tips as previously described (61). Briefly, root tips of *E. dens-canis* and *E. sibiricum* were harvested from seedlings and bulbs of cultivated plants, respectively, pre-treated with ice-cold water for 16 h, fixed in ethanol/acetic acid (3:1) fixative for 24 h at 4°C and stored at -20°C until further use. Selected root tips were rinsed in distilled water (twice for 5 min) and citrate buffer (10 mM sodium citrate, pH 4.8; twice for 5 min), and digested in 0.3% cellulase, cytohelicase and

pectolyase (all Sigma-Aldrich, St. Louis, MO, USA) in citrate buffer at 37°C for 3 h. After digestion, individual root tips were dissected on a microscope slide in 20 µl acetic acid and spread on the slide placed on a metal hot plate (50°C) for c. 30 s. Then, the preparation was fixed in freshly prepared ethanol/acetic acid (3:1) fixative by dropping the fixative around the drop of acetic acid and into it. The preparation was dried using a hair dryer and staged using a phase contrast microscope. Chromosome preparations were treated with 100 µg/ml RNase in 2× sodium saline citrate (SSC; 20× SSC: 3 M sodium chloride, 300 mM trisodium citrate, pH 7.0) for 60 min and with 0.1 mg/ml pepsin in 0.01 M HCl at 37°C for 5 min; then postfixed in 4% formaldehyde in 2× SSC for 10 min, washed in 2× SSC twice for 5 min, and dehydrated in an ethanol series (70%, 90%, and 100%, 2 min each).

Fluorescence *in situ* Hybridization. The 821 bp ETS from *Erythronium* and 220 bp 26S rDNA from *Nicotiana* (62) were used for *in situ* localization of rDNA loci in *E. dens-canis* and *E. sibiricum*. The two rDNA probes were labelled with biotin-dUTP or digoxigenin-dUTP by nick translation as described in (63). 200 ng of each rDNA probe were pooled together, ethanol precipitated, dissolved in a 20 µl mixture containing 50% formamide, 10% dextran sulfate and 2× SSC, and pipetted onto microscopic slides. The slides were heated at 80°C for 2 min and incubated at 37°C overnight. Hybridized probes were visualized through fluorescently-labelled antibodies against biotin-dUTP and digoxigenin-dUTP as in (63). Chromosomes were counterstained with 4′,6-diamidino-2-phenylindole (DAPI, 2 µg/ml) in Vectashield antifade. Fluorescence signals were analyzed and photographed using a Zeiss Axioimager epifluorescence microscope and a CoolCube camera (MetaSystems, Altlussheim, Germany). Individual images were merged and processed using the Photoshop CS software (Adobe Systems, San Jose, CA, USA).

Construction and Screening of a BAC Library from *E. dens-canis*. HMW DNA was prepared according to a protocol modified from that in ref (64). The method takes advantage of purification and sorting of nuclei using flow cytometry and embedding them into agarose plugs. Briefly, 30 g of fresh leaf tissue was fixed in 2% formaldehyde for 20 minutes and homogenized. Nuclei were concentrated from the resulting suspension by passing it through a FACSAria device (Becton Dickinson, San José, CA, USA) and were embedded in 20 µl agarose plugs (each plug harbored around 2 x 105 nuclei). The plugs were solidified on ice and incubated in lysis buffer (0.5 M EDTA pH 9.0, 1% lauroylsarcosine, 0.2 mg/ml proteinase K) at 37°C for 24 hours twice to remove residual proteins. HindIII restriction fragments of the obtained HMW DNA were size-selected by pulse field gel electrophoresis (PFGE) and ligated into pAGIBAC vector. *E. coli* DH10B-T1R strain was used as host for the cloning. Unique clones (173 610 in number and having average insert size of 115 kbp) were arranged in pools of 90 and tested. Primers for dicot and monocot ITS2 (Table S6) were used to screen the low-coverage (~0.81×) library by quantitative PCR (qPCR). Colony picking was carried out on the positive pools by a high throughput automated station QPix2 XT (Molecular Devices, San José, CA, USA). Individual BAC clones were validated by qPCR. Insert size was estimated using NotI enzymatic digestion and PFGE electrophoresis. BAC-ends sequences (BESs) were performed using Sanger sequencing. Construction and screening of the BAC library as well as its long-term storage (under the name of Ede-B-3544H) was achieved at the Plant Genomic Center (INRAE-CNRGV), Toulouse (France).

BAC Clone Sequence Assembly and Analysis. Upon screening, multiple BAC clones tested positive for the monocot marker and a single clone (1J24) was found to contain the dicot marker. BAC clones 1J24 and monocot marker-specific 18I01 were tagged and pooled before performing sequencing on a PacBio Sequel II system with the CCS mode (chemistry 3.0). Raw reads were corrected and demultiplexed. The HiFi reads were then filtered by eliminating the residual *E. coli* and vector sequences and assembled with Canu-1.9 (1J24) and Hifiasm-0.12 (18I01). A quality control of the assembly was conducted. First, BESs obtained by the Sanger method were aligned to the ends of the assembled contigs. Second, the coverage of assemblies and the length of the obtained contigs were checked. The assemblies of BAC clones 1J24 and 18I01 had 593× and 976× coverages and 125 kbp and 90 kbp contig lengths, respectively. Each assembly resulted in one single contig. Annotation and physical map creation of obtained BAC clone sequences was performed in Geneious Prime v.2019.2.1.

11

Searches for conserved domains within BAC clone sequences were performed at the National Center for Biotechnology Information website (https://www.ncbi.nlm.nih.gov/Structure/cdd/wrpsb.cgi). The online version of *LTR_FINDER* (65) was used to check for retrotransposon LTRs. BAC clones were further explored for transposon-like sequences with Censor (https://www.girinst.org/censor/index.php) (66) by restricting the search to *Viridiplantae*. Hits not containing simple repeats and those of at least 150 bp in length were considered.

Test for Dicot rDNA Transcription in *Erythronium* and Assessing the Relative Fold Difference Between the Two rDNA Types. In order to test for the transcription of dicot-type rDNA in *Erythronium*, total RNA was isolated with the RNEasy Plant Mini Kit (Qiagen, Hilden, Germany) from actively growing (in October) roots of an *E. dens-canis* bulb. RNA was treated with DNase, and processed for cDNA synthesis with RevertAid First Strand cDNA Synthesis Kit (Thermo Scientific, Vilnius, Lithuania). Complementary DNA and gDNA (for control) were then used as templates in the upcoming regular PCR using primers for dicot and monocot ITS2 (Table S6). A SQ-PCR approach was employed using Maxima SYBR Green qPCR Master Mix (Thermo Scientific, Waltham, MA, USA) and primers for dicot and monocot ITS2. The experiment was extended with gDNAs of one *E. caucasicum* sample and four additional *E. dens-canis* samples from distant localities of the species' range (Table S1). Each sample-primer pair combination was run in triplicate. Raw fluorescence data provided by a Rotor-Gene 6000 instrument (Corbett Research, Mortlake, Australia) were analyzed in the LinRegPCR v.11.0 software. Relative quantification was achieved with the RFD = $E^{-delta\ Cq}$ formula. Aliquots of the completed SQ-PCR reactions were ultimately verified by agarose-gel electrophoresis and ethidium bromide-UV detection.

Nanopore Long-Read Sequencing and Data Analysis. Oxford Nanopore Technologies' (ONT) long-read sequencing service was ordered from Deep Seq (University of Nottingham, Nottingham, UK). Bulb-extracted gDNA of ~25 kbp in size was used as input for the preparation of the standard SQK-LSK109 library as per ONT's protocol (ONT, Oxford, UK). The library was run over one PromethION Flow Cell. Base calling was completed in PromethION's sequencing software MinKnow by using the default quality threshold of 7. Resulting 12.6 million reads (~68.2 Gigabases (Gb) of data) had N50 read length of 9.9 kbp. A total of 64998 contigs with minimum, maximum and average lengths of 1 kbp, 1.38 Mbp and 5.69 kbp, respectively were assembled using CANU v 1.8 (67) from the raw Nanopore data. Polishing of the CANU assembly was attempted using Pilon (68) (https://github.com/broadinstitute/pilon) with default parameters, using a preprocessed paired-end Illumina dataset (see below). Both the raw reads and the contigs were screened for dicot/monocot ITS1/ITS2 fragments (Table S3) in Geneious Prime. In terms of the number and the length of Nanopore reads containing *Potentilla*-specific rDNA sequences, the assembly did not turn out to be superior to the original read dataset (Table S4). Raw Nanopore sequences that were analyzed and presented in the current study were included in the supplementary dataset (Datasets S1 and S2).

Illumina Whole Genome Shotgun (WGS) Sequencing and Read Analysis.
A TruSeq shotgun library (of 350 bp insert-size) was prepared from gDNA of *E. dens-canis* and paired-end (150 bp) sequenced via a HiSeqX platform at Macrogen B.V. Fifty gigabases of raw reads were preprocessed to remove low-quality reads and controls and to trim adapter sequences by using a combination of tools provided in the BBMap project (69, 70). A *de novo* assembly of Illumina reads was performed in SPAdes (71). *Potentilla*-specific ITS and partial-ETS sequences were extracted from contigs following a screen in Geneious Prime with dicot ITS1/ITS2 sequences.

In case of *E. caucasicum* 5 Gb of WGS data was obtained. A shotgun gDNA library of 350 bp insert size was prepared and paired-end (150 bp) sequenced at Novogene (Cambridge, UK) using the Novaseq6000 platform. Raw reads were imported into Geneious Prime, paired, and end-trimmed using the error probability limit of 0.05. Paired reads were then mapped to dicot ITS1/ITS2 reference sequences. Mapped reads were assembled into contigs and consensuses of the contigs were generated by using the built-in Geneious assembler.

12

Acknowledgments

13

References

1. E. Aubin, M. El Baidouri, O. Panaud, Horizontal gene transfers in plants. *Life* **11**, 857 (2021).
2. J. Van Etten, D. Bhattacharya, Horizontal gene transfer in eukaryotes: Not if, but how much? *Trends Genet.* **36**, 915-925 (2020).
3. F.-W. Li *et al.*, Horizontal transfer of an adaptive chimeric photoreceptor from bryophytes to ferns. *Proc. Natl. Acad. Sci. U.S.A.* **111**, 6672-6677 (2014).
4. J. Xia *et al.*, Whitefly hijacks a plant detoxification gene that neutralizes plant toxins. *Cell* **184**, 1693-1705.e1617 (2021).
5. C. Phansopa, L. T. Dunning, J. D. Reid, P.-A. Christin, Lateral gene transfer acts as an evolutionary shortcut to efficient C4 biochemistry. *Mol. Biol. Evol.* **37**, 3094-3104 (2020).
6. C. C. Davis, Z. Xi, Horizontal gene transfer in parasitic plants. *Curr. Opin. Plant Biol.* **26**, 14-19 (2015).
7. D. A. Wickell, F.-W. Li, On the evolutionary significance of horizontal gene transfers in plants. *New Phytol.* **225**, 113-117 (2020).
8. S. G. S. Hibdige, P. Raimondeau, P.-A. Christin, L. T. Dunning, Widespread lateral gene transfer among grasses. *New Phytol.* **230**, 2474-2486 (2021).
9. M. El Baidouri *et al.*, Widespread and frequent horizontal transfers of transposable elements in plants. *Genome Res.* **24**, 831-838 (2014).
10. X. Diao, M. Freeling, D. Lisch, Horizontal transfer of a plant transposon. *PLoS Biol.* **4**, e5 (2006).
11. M. Park, P.-A. Christin, J. L. Bennetzen, Sample sequence analysis uncovers recurrent horizontal transfers of transposable elements among grasses. *Mol. Biol. Evol.* **38**, 3664-3675 (2021).
12. A. Roulin, B. Piegu, R. A. Wing, O. Panaud, Evidence of multiple horizontal transfers of the long terminal repeat retrotransposon *RIRE1* within the genus *Oryza*. *Plant J.* **53**, 950-959 (2008).
13. F. Martinez-Seidel, O. Beine-Golovchuk, Y.-C. Hsieh, J. Kopka, Systematic review of plant ribosome heterogeneity and specialization. *Front. Plant Sci.* **11** (2020).
14. S. Wicke, A. Costa, J. Muñoz, D. Quandt, Restless 5S: The re-arrangement(s) and evolution of the nuclear ribosomal DNA in land plants. *Mol. Phylogen. Evol.* **61**, 321-332 (2011).
15. B. G. Baldwin *et al.*, The ITS region of nuclear ribosomal DNA: A valuable source of evidence on angiosperm phylogeny. *Ann. Mo. Bot. Gard.* **82**, 247-277 (1995).
16. A. Kovařík *et al.*, Concerted evolution of 18–5.8–26S rDNA repeats in *Nicotiana* allotetraploids. *Biol. J. Linn. Soc.* **82**, 615-625 (2004).
17. W. Wang *et al.*, Astonishing 35S rDNA diversity in the gymnosperm species *Cycas revoluta* Thunb. *Chromosoma* **125**, 683-699 (2016).
18. L.-Q. Xiao, M. Möller, H. Zhu, High nrDNA ITS polymorphism in the ancient extant seed plant *Cycas*: Incomplete concerted evolution and the origin of pseudogenes. *Mol. Phylogen. Evol.* **55**, 168-177 (2010).
19. A. Kovařík *et al.*, Rapid concerted evolution of nuclear ribosomal DNA in two *Tragopogon* allopolyploids of recent and recurrent origin. *Genetics* **169**, 931-944 (2005).
20. J. Sochorová *et al.*, Gene conversion events and variable degree of homogenization of rDNA loci in cultivars of *Brassica napus*. *Ann. Bot.* **119**, 13-26 (2016).
21. R. Jain, M. C. Rivera, J. A. Lake, Horizontal gene transfer among genomes: The complexity hypothesis. *Proc. Natl. Acad. Sci. U.S.A.* **96**, 3801-3806 (1999).
22. K. Kitahara, K. Miyazaki, Revisiting bacterial phylogeny: Natural and experimental evidence for horizontal gene transfer of 16S rRNA. *Mob. Genet. Elements* **3**, e24210 (2013).
23. M. Sato, K. Miyazaki, Phylogenetic network analysis revealed the occurrence of horizontal gene transfer of 16S rRNA in the genus *Enterobacter*. *Front. microbiol.* **8** (2017).
24. R.-M. Tian, L. Cai, W.-P. Zhang, H.-L. Cao, P.-Y. Qian, Rare events of intragenus and intraspecies horizontal transfer of the 16S rRNA gene. *Genome Biol. Evol.* **7**, 2310-2320 (2015).

25. A. Yabuki, T. Toyofuku, K. Takishita, Lateral transfer of eukaryotic ribosomal RNA genes: an emerging concern for molecular ecology of microbial eukaryotes. *ISME J.* **8**, 1544-1547 (2014).

26. J. Xie *et al.*, Intergeneric transfer of ribosomal genes between two fungi. *BMC Evol. Biol.* **8**, 87 (2008).

27. V. Mahelka *et al.*, Multiple horizontal transfers of nuclear ribosomal genes between phylogenetically distinct grass lineages. *Proc. Natl. Acad. Sci. U.S.A.* **114**, 1726-1731 (2017).

28. D. Rungger, M. Crippa, The primary ribosomal DNA transcript in eukaryotes. *Prog. Biophys. Mol. Biol.* **31**, 247-269 (1978).

29. G. A. Allen, D. E. Soltis, P. S. Soltis, Phylogeny and biogeography of *Erythronium* (Liliaceae) inferred from chloroplast matK and nuclear rDNA ITS sequences. *Syst. Bot.* **28**, 512-523 (2003).

30. J. C. B. Clennett, M. W. Chase, F. Forest, O. Maurin, P. Wilkin, Phylogenetic systematics of *Erythronium* (Liliaceae): morphological and molecular analyses. *Bot. J. Linn. Soc.* **170**, 504-528 (2012).

31. L. Bartha, N. V. Stepanov, J. Rukšāns, H. L. Banciu, L. Keresztes, Non-monophyly of Siberian *Erythronium* (Liliaceae) leads to the recognition of the formerly neglected *Erythronium sajanense*. *J. Plant Res.* **128**, 721-729 (2015).

32. S. Kumar, G. Stecher, M. Suleski, S. B. Hedges, TimeTree: A resource for timelines, timetrees, and tivergence times. *Mol. Biol. Evol.* **34**, 1812-1819 (2017).

33. S. Siljak-Yakovlev *et al.*, Towards a genome size and chromosome number database of Balkan Flora: C-values in 343 taxa with novel values for 242. *Adv. Sci. Lett.* **3**, 190-213 (2010).

34. J. Pellicer, O. Hidalgo, S. Dodsworth, I. J. Leitch, Genome size diversity and its impact on the evolution of land plants. *Genes* **9**, 88 (2018).

35. L. Bartha *et al.*, Patterns of plastid DNA differentiation in *Erythronium* (Liliaceae) are consistent with allopatric lineage divergence in Europe across longitude and latitude. *Plant Syst. Evol.* **301**, 1747-1758 (2015).

36. N. L. Persson, I. Toresen, H. L. Andersen, J. E. E. Smedmark, T. Eriksson, Detecting destabilizing species in the phylogenetic backbone of *Potentilla* (Rosaceae) using low-copy nuclear markers. *AoB PLANTS* **12** (2020).

37. A. A. Kechaykin, A. I. Shmakov, T. Gregor, J. Paule, D. A. German, (2820) Proposal to conserve the name *Potentilla* (Rosaceae: Potentilleae) with a conserved type. *TAXON* **70**, 680-681 (2021).

38. C. Dobeš, J. Paule, A comprehensive chloroplast DNA-based phylogeny of the genus *Potentilla* (Rosaceae): Implications for its geographic origin, phylogeography and generic circumscription. *Mol. Phylogen. Evol.* **56**, 156-175 (2010).

39. Z. Yang *et al.*, Convergent horizontal gene transfer and cross-talk of mobile nucleic acids in parasitic plants. *Nat. Plants* **5**, 991-1001 (2019).

40. G. Bourque *et al.*, Ten things you should know about transposable elements. *Genome Biol.* **19**, 199 (2018).

41. L. T. Dunning *et al.*, Lateral transfers of large DNA fragments spread functional genes among grasses. *Proc. Natl. Acad. Sci. U.S.A.* **116**, 4416-4425 (2019).

42. L. A. Graham, P. L. Davies, Horizontal gene transfer in vertebrates: A fishy tale. *Trends Genet.* **37**, 501-503 (2021).

43. V. Mahelka *et al.*, A *Panicum*-derived chromosomal segment captured by *Hordeum* a few million years ago preserves a set of stress-related genes. *Plant J.* **105**, 1141-1164 (2021).

44. T. Kado, H. Innan, Horizontal gene transfer in five parasite plant species in Orobanchaceae. *Genome Biol. Evol.* **10**, 3196-3210 (2018).

45. C. C. Davis, W. R. Anderson, K. J. Wurdack, Gene transfer from a parasitic flowering plant to a fern. *Proc. R. Soc. B: Biol. Sci.* **272**, 2237-2242 (2005).

46. B. Dhillon *et al.*, Horizontal gene transfer and gene dosage drives adaptation to wood colonization in a tree pathogen. *Proc. Natl. Acad. Sci. U.S.A.* **112**, 3451-3456 (2015).

47. S. Yoshida, S. Maruyama, H. Nozaki, K. Shirasu, Horizontal gene transfer by the parasitic plant *Striga hermonthica*. *Science* **328**, 1128-1128 (2010).

48. T. Fukuda, S. Nakamura, Biotic interaction between a rust fungus, *Uromyces erythronii* Pass., and its host plant, *Erythronium japonicum* Decne. (Liliaceae). *Plant Species Biol.* **2**, 75-83 (1987).

49. T. Nagy, W. P. Pfliegler, A. Takács, J. Tökölyi, A. Molnár V., Distribution, infection rates and DNA barcoding of *Uromyces erythronii* (Pucciniaceae), a parasite of *Erythronium* (Liliaceae) in Europe. *Willdenowia* **49**, 13-20, 18 (2019).

50. S. Helfer, Overview of the rust fungi (Uredinales) occurring on Rosaceae in Europe. *Nova Hedwigia* **81**, 325-370 (2005).

51. A. P. Ramos *et al.*, Flow cytometry reveals that the rust fungus, *Uromyces bidentis* (Pucciniales), possesses the largest fungal genome reported—2489 Mbp. *Mol. Plant Pathol.* **16**, 1006-1010 (2015).

52. S. Tavares *et al.*, Genome size analyses of Pucciniales reveal the largest fungal genomes. *Front. Plant Sci.* **5** (2014).

53. S. Lengyel, A. D. Gove, A. M. Latimer, J. D. Majer, R. R. Dunn, Convergent evolution of seed dispersal by ants, and phylogeny and biogeography in flowering plants: A global survey. *Perspect. Plant Ecol. Evol. Syst.* **12**, 43-55 (2010).

54. J. Huang, Horizontal gene transfer in eukaryotes: The weak-link model. *Bioessays* **35**, 868-875 (2013).

55. P.-A. Christin *et al.*, Adaptive evolution of C4 photosynthesis through recurrent lateral gene transfer. *Curr. Biol.* **22**, 445-449 (2012).

56. S. Asker, Some viewpoints on *Fragaria* × *Potentilla* intergeneric hybridization. *Hereditas* **67**, 191-190 (1971).

57. F. Pustahija, N. Bašić, S. Siljak-Yakovlev, "Chromosome numbers and genome size data for some Balkan species" in Flora Mediterranea 28*,* G. Kamari, C. Blanché, S. Siljak-Yakovlev, Eds. (International Foundation pro Herbario Mediterraneo, Palermo, 2018), vol. 28, chap. Mediterranean plant karyological data – 28, pp. 420-428.

58. T. J. White, T. Bruns, S. Lee, J. Taylor, "Amplification and direct sequencing of fungal ribosomal RNA genes for phylogenetics" in PCR Protocols: A Guide to Methods and Applications*,* M. A. Innis, D. H. Gelfand, J. J. Shinsky, T. J. White, Eds. (Academic Press, San Diego, 1990), pp. 315-322.

59. T. Feng *et al.*, Phylogenetic study of the tribe Potentilleae (Rosaceae), with further insight into the disintegration of *Sibbaldia. J. Syst. Evol.* **55**, 177-191 (2017).

60. L.-T. Nguyen, H. A. Schmidt, A. von Haeseler, B. Q. Minh, IQ-TREE: A fast and effective stochastic algorithm for estimating maximum-likelihood phylogenies. *Mol. Biol. Evol.* **32**, 268-274 (2014).

61. T. Mandáková, M. A. Lysak, Chromosome preparation for cytogenetic analyses in *Arabidopsis. Curr. Protoc. Plant Biol.* **1**, 43-51 (2016).

62. K. Y. Lim *et al.*, Gene conversion of ribosomal DNA in *Nicotiana tabacum* is associated with undermethylated, decondensed and probably active gene units. *Chromosoma* **109**, 161-172 (2000).

63. T. Mandáková, M. A. Lysak, Painting of *Arabidopsis* chromosomes with chromosome-specific BAC clones. *Curr. Protoc. Plant Biol.* **1**, 359-371 (2016).

64. J. Šafář *et al.*, Creation of a BAC resource to study the structure and evolution of the banana (*Musa balbisiana*) genome. *Genome* **47**, 1182-1191 (2004).

65. Z. Xu, H. Wang, LTR_FINDER: an efficient tool for the prediction of full-length LTR retrotransposons. *Nucleic Acids Res.* **35**, W265-268 (2007).

66. O. Kohany, A. J. Gentles, L. Hankus, J. Jurka, Annotation, submission and screening of repetitive elements in Repbase: RepbaseSubmitter and Censor. *BMC Bioinform.* **7**, 474 (2006).

67. S. Koren *et al.*, Canu: scalable and accurate long-read assembly via adaptive *k*-mer weighting and repeat separation. *Genome Res.* **27**, 722-736 (2017).

68. B. J. Walker *et al.*, Pilon: an integrated tool for comprehensive microbial variant detection and genome assembly improvement. *PLoS ONE* **9**, e112963 (2014).

69. B. Bushnell (2016) BBMap short read aligner. (Univ. California, Berkeley, California.).

70. B. Bushnell, J. Rood, E. Singer, BBMerge – Accurate paired shotgun read merging via overlap. *PLoS ONE* **12**, e0185056 (2017).

71. A. Bankevich *et al.*, SPAdes: a new genome assembly algorithm and its applications to single-cell sequencing. *J. Comput. Biol.* **19**, 455-477 (2012).

72. B. Q. Minh, M. A. T. Nguyen, A. von Haeseler, Ultrafast approximation for phylogenetic Bootstrap. *Mol. Biol. Evol.* **30**, 1188-1195 (2013).
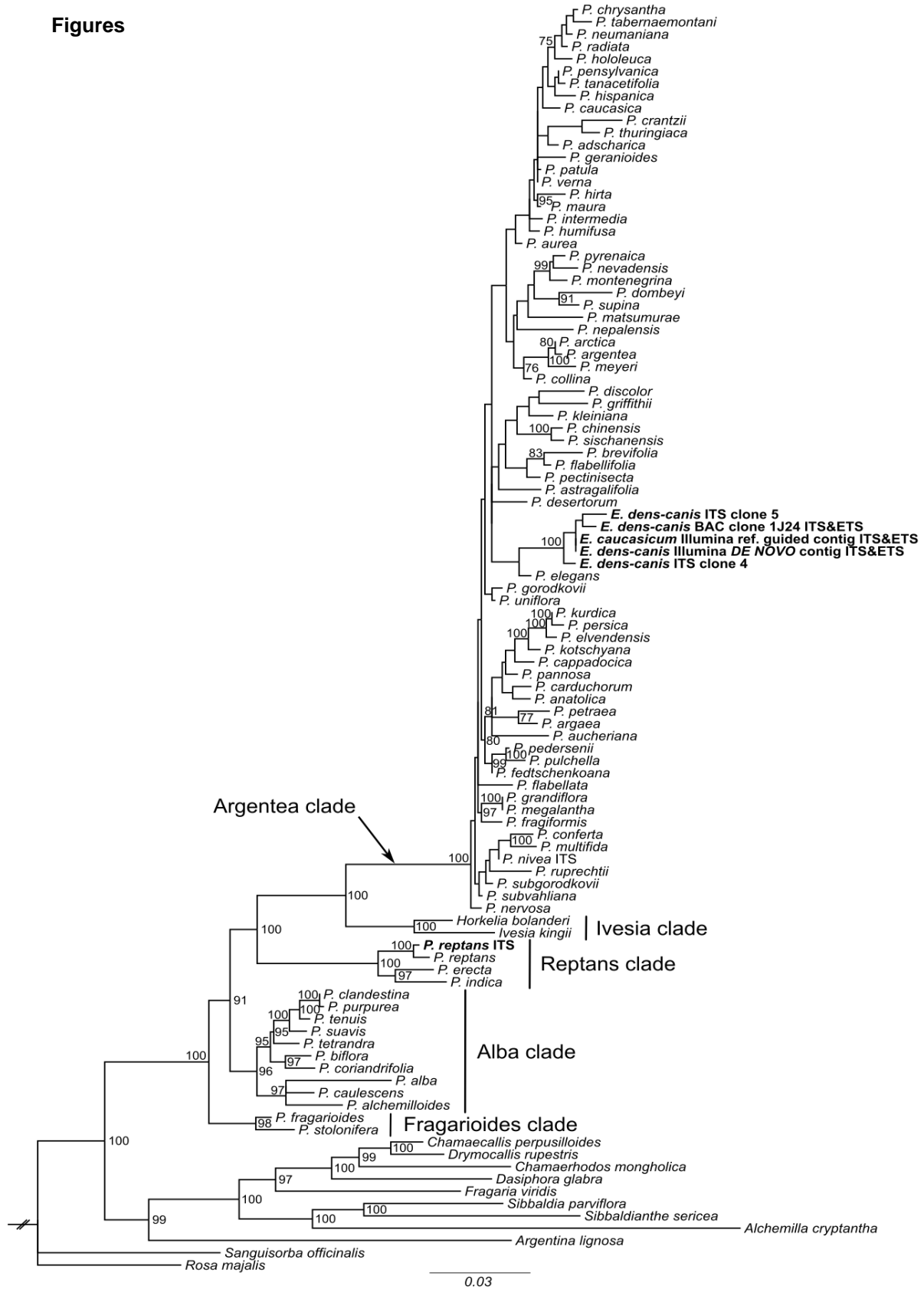
Figures

Figure 1. IQ-TREE ML phylogram based on *Potentilla* and *Potentilla*-specific *Erythronium* ITS and partial ETS sequences. Terminal names with newly generated sequences are in bold. Nodal support percentages (not shown below 75) were derived from 1000 ultrafast bootstrap [UFBoot (72)] replicates.
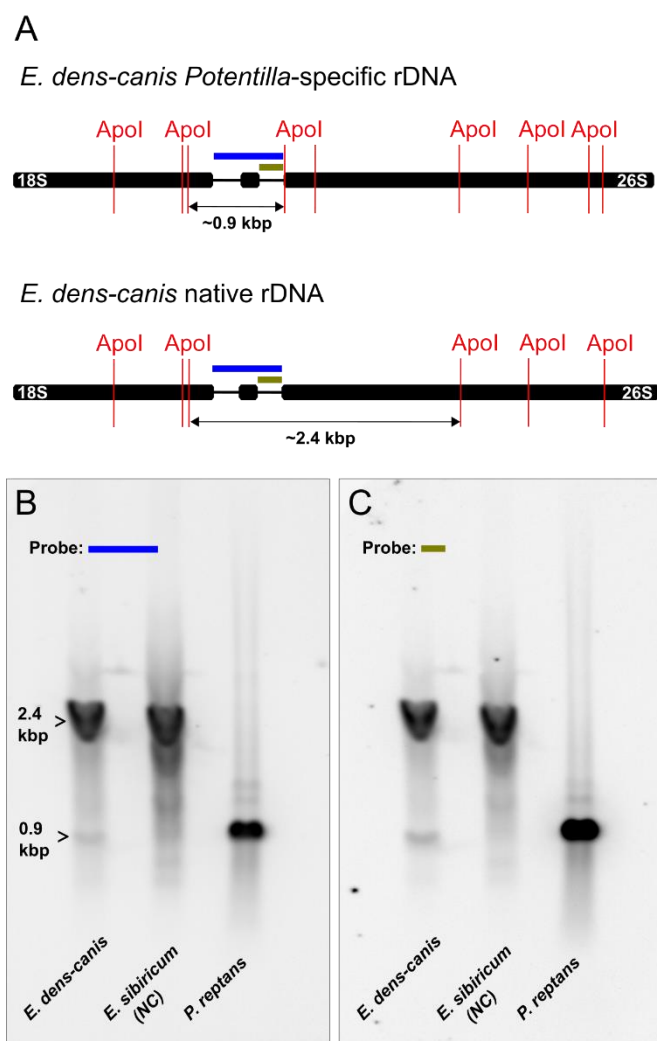
Figure 2. Southern blot analysis of rDNA fragments produced by *Apo*I-HF from gDNA of *E. dens-canis*, *E. sibiricum* and *P. reptans*. (A). Restriction map of *Apo*I-HF targeting the native and alien rDNA of *E. dens-canis*. Probes consisted of *P. reptans* ITS region (B) and *Potentilla*-type ITS2 cloned from *E. dens-canis* (C). Arrows indicate *Potentilla*- (~0.9 kbp) and *Erythronium*-specific (~ 2.4 kbp) rDNA fragments. NC, negative control.
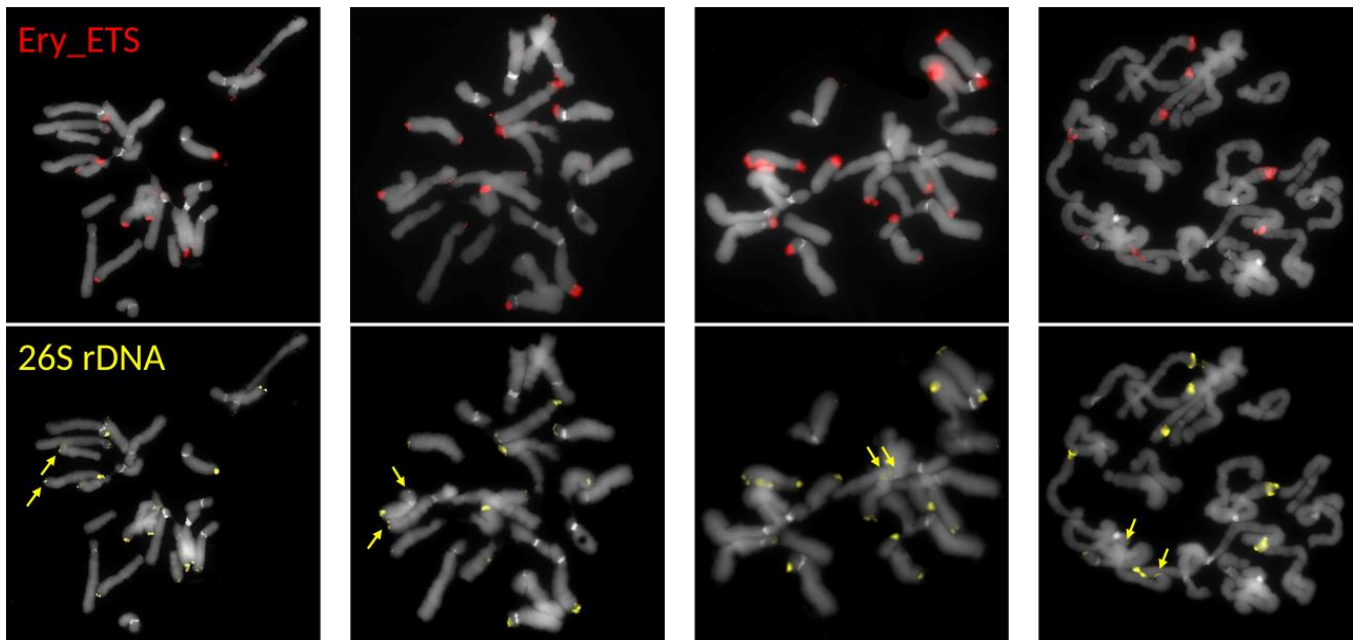
Figure 3. Chromosomal localization of rDNA in *Erythronium dens-canis.* Four mitotic chromosome complements were hybridized with *Erythronium* ETS (Ery_ETS, red signals) and *Nicotiana* 26S rDNA probe (yellow signals). Extraneous sites identified by the 26S probe (marked by arrows) are likely representing the sites of alien rDNA. Chromosomes were counterstained with DAPI.
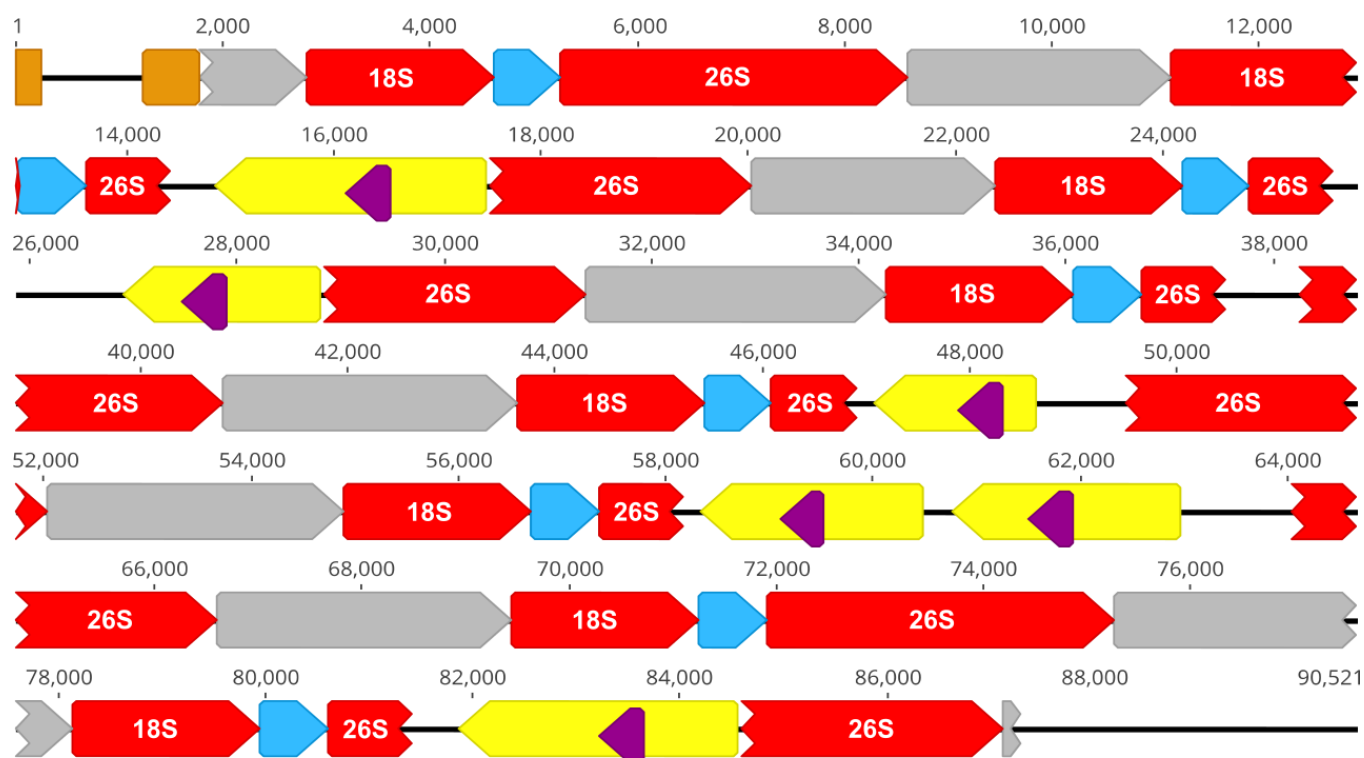
Figure 4. Physical map of *E. dens-canis* BAC clone 1J24 containing rDNA of exclusively *Potentilla*-origin.
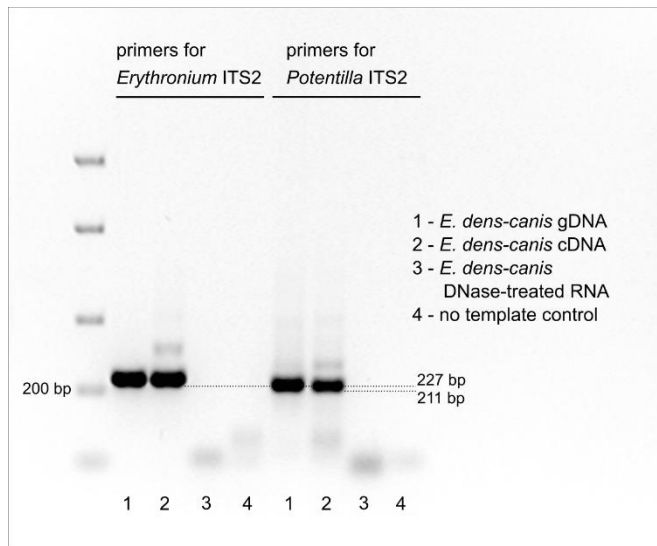
Figure 5. Evidence of transcription of alien rDNA in *E. dens-canis* based on the amplification of *Potentilla*-specific ITS2 from cDNA. PCR products were visualized according to ethidium bromide-UV detection and were based on the templates and primers indicated in the body of figure. ITS2 fragment sizes include lengths of primers.