Emilie Aubin [a+], Christel Llauro [ab+], Joseph Garrigue [c], Marie Mirouze [ad], Olivier Panaud [ae*], Moaine El Baidouri [ab*]

Genome-Wide Analysis of Horizontal Transfer in Non-Model Wild Species from a Natural Ecosystem Reveals New Insights into Genetic Exchange in Plants.

a. Université de Perpignan Via Domitia, Laboratoire Génome et Développement des Plantes, UMR5096, 66860 Perpignan, France

b. Centre National de la Recherche Scientifique, Laboratoire Génome et Développement des Plantes, UMR5096, 66860 Perpignan, France

c. Réserve Naturelle Nationale de la forêt de la Massane, France

d. IRD, UMR232 DIADE, 34394 Montpellier, France

e. Institut Universitaire de France, 75231 Paris cedex 05, France

Corresponding authors:

Moaine El Baidouri (moaine.elbaidouri@univ-perp.fr)

Olivier Panaud (panaud@univ-perp.fr)

+ : These authors contributed equally.

Abstract

Background

Horizontal transfer (HT) refers to the exchange of genetic material between divergent species by mechanisms other than reproduction. In recent years, several studies have demonstrated HTs in eukaryotes, particularly in the context of parasitic relationships and in model species. However, very little is known about HT in natural ecosystems, especially those involving non-parasitic wild species, and the nature of the ecological relationships that promote these HTs.

Results

In this work, we conducted a pilot study investigating HTs by sequencing the genomes of 17 wild non-model species from a natural ecosystem, the Massane forest, located in southern France. To this end, we developed a new computational pipeline called INTERCHANGE that

is able to characterize HTs at the whole genome level without prior annotation and directly in the raw sequencing reads. Using this pipeline, we identified 12 HT events, half of which occurred between lianas and trees. We found that mainly low copy number LTRs retrotransposons from the Copia superfamily were transferred between these wild plant species, especially those of the Ivan and Ale lineages.

Conclusion

This study revealed a possible new route for HTs between non-parasitic plants and provides new insights into the genomic characteristics of horizontally transferred DNA in plant genomes.

Keywords

Horizontal Transfer, Natural Ecosystem, Climbing Plants, Trees, LTRs-retrotransposons, Genome Evolution, Genomics

Background

Horizontal transfer (HT) is a process by which genetic material is exchanged between two distinct species without reproduction. HTs are well documented in prokaryotes and considered to play a major role in the adaptation and colonization of new ecological niches [1]. The rapid spread of antibiotic resistance genes among bacteria is a good example of the adaptive role of HTs [2]. Although HTs are thought to be less common in eukaryotes, numerous examples of HTs between multicellular eukaryotes such as plants, animals and insects have been reported in recent years [3–8]. Indeed, over the past decades, the number of sequenced and assembled genomes has steadily increased, facilitating the discovery of several horizontally transferred genes and transposable elements (TEs) between eukaryotes [9–12]. Some of these described HTs were adaptive [13–18]. For instance, in plants, there are several major cases of HTs leading to adaptive innovations such as the recent case of a detoxification gene transmitted horizontally from an endophytic fungus to a wild cereal (*Thinopyrum*) allowing the latter to become resistant to *Fusarium* [17]. This naturally transferred gene was further introduced by breeders into wheat through wide hybridization, resulting in broad resistance to ear blight and crown rot. Another recent example is the acquisition of a detoxification gene through HT in the whitefly, a plant feeding insect, which enables it to overcome host plant defences [18].

Host-parasite interactions has been shown to promote HTs in eukaryotes, such as in parasitic plants. This is particularly true for the *Orobanchaceae* [19], *Striga* [20], *Cuscuta* [21] and *Rafflesia* [22,23]. In animals, a bloodsucking insect has for instance transferred multiple transposon families between different mammalian species it parasitizes [24]. Natural grafting [25,26] could also facilitate the occurrence of HTs between plant species. There is also strong evidence that HTs occurs between species that do not share any host/parasite interaction [9–12,27].

The vast majority of previous studies on HTs have been conducted using genomic data from public databases of model species for which sympatric relationships and the nature of biological interactions are not always known. This represents a major hurdle in attempts to understand the mechanisms and nature of biotic relationships that can promote HTs in natural ecosystems. Furthermore, while previous reports have shown that both genes and TEs can be horizontally transferred between eukaryotes such as plants, it is not clear whether these two genomic components transfer at the same rate or whether certain types of genes or TEs are more frequently transferred than others. This is because these studies have focused on a group of specific genes or TEs, mainly because the methodologies used so far require prior annotation of the sequences of interest (genes or TEs) limiting any investigation of HTs at the whole genome level.

To address these questions, we conducted a pilot study aiming to investigate HTs in wild plant species from a natural ecosystem, the Massane beech forest located in southern France, considered as one of the last relict forests of the Quaternary Period in Europe [28]. We sampled 17 wild non-model species from this reservation, including trees, climbing plants, herbaceous species and fungi. Through de novo whole-genome sequencing of these species using *Illumina* and the development of a new computational pipeline named INTERCHANGE (for horIzoNtal TransfER CHAracterization in Non-assembled Genome) 12 HTs involving 8 species have been identified. These HTs involve TEs, specifically low copy number LTR-retrotransposons from the *Copia* superfamily. Furthermore, we found that some climbing plants underwent multiple HT events with tree species which could constitute a new route of HT between non-parasitic plants.

Results

INTERCHANGE a new strategy for horizontal transfer identification at the whole-genome scale using short read sequencing data

The inference of HTs is usually based on the use of three criteria [29,30]: i) high sequence similarity between evolutionary divergent species (HS); ii) phylogenetic incongruence between the evolutionary history of the species and that of the transferred sequence (PI); iii) patchy distribution of the transferred sequence in the phylogeny of the species (PD). In other words, no homologous sequence of the transferred DNA is found in the closest relative of the recipient species. Previous methods used for HT detection based on one or two of the above criteria require the prior genome assembly and annotation. This makes HT studies between wild non-model species for which there is no reference genome or annotation available very challenging. We have therefore developed a new pipeline to identify highly conserved regions that may arise from HT between two or several genomes of non-model species using raw short-read sequencing data. Briefly, this automatic pipeline that we called INTERCHANGE, (i) first identifies similar reads derived from conserved locus between the studied species using a K-mer approach (ii) assembles these reads into scaffolds (iii) annotate the scaffolds (iv) and test for high sequence similarity (HS) by comparing the sequence identity between conserved scaffolds with that of orthologous genes. Those HT candidates are then manually tested for the PI and PD criteria. The main steps of this pipeline are shown in Figure 1 and described in details in the method section.

INTERCHANGE validation using simulated horizontal transfers

To test the accuracy of INTERCHANGE, we simulated HT events involving both genes and transposable elements (TEs) of various classes and superfamilies between three plant models: *Arabidopsis thaliana*, *Oryza sativa* and *Brachypodium distachyon* (Figure 2, see method for details). These species were chosen because they have high-quality, well-annotated reference genomes and present a contrasting evolutionary divergence times. *A. thaliana* and *O. sativa* share a last common ancestor 160 million years ago, while the two Poaceae species, O sativa and B. distachyon, diverged from each other less than 46 million years ago (http://www.timetree.org/). HTs were simulated between these species with contrasting divergence times in both directions, using each species as a donor and a recipient (Figure 2-a). Using the assembled reference genomes, we inserted 100 genes and 100 TEs from *A. thaliana* into the *O. sativa* reference genome, and 100 genes and 100 TEs from *O. sativa* into the *A. thaliana* reference genome, resulting in 400 simulated HT events between these two species (see methods section). Before inserting these sequences into the recipient genomes, random artificial mutations were introduced to simulate both ancient and recent HTs as described in

the method section. Next, *Illumina* short reads were simulated for each genome harboring these *in silico* HTs. The same HT simulation was carried out between *O. sativa* and *B. distachyon*.

The INTERCHANGE pipeline was applied to these species using simulated short read sequencing data from the artificially modified genomes that incorporated simulated HT events. INTERCHANGE was able to identify 90% (361 out of 400) of simulated HT events between *A. thaliana* and *O. sativa* (Figure 2-b) (HS = 80%). 10% of simulated HTs were false negative (39 out of 400) and only one single scaffold candidate was a false positive. Among the identified HTs, 52 % were genes and 48 % were TEs. For the pair *O. sativa* and *B. distachyon*, INTERCHANGE identified 73% (292/400) of simulated HTs (53% genes and 47% TEs) with a false negative rate of 27% (109 out of 400) and 743 false positives. 85% of these false positives correspond to scaffolds smaller than 500 bp and correspond to highly conserved genes with a sequence similarity higher than the HS threshold (i.e HS = 87%). False negative correspond to simulated HTs with lower sequence similarity than the threshold. This increase in false negative rate and false positive rate compared to *A. thaliana* and *O. sativa* can be explained by the smaller evolutionary distance between the two Poaceae species. As shown in Figure 2-b, when only scaffolds with longer size are considered however, the number of false positives decreases significantly, with a smaller decrease in true positives and without affecting the relative proportion of detected genes and TEs. For instance, by limiting the candidates to those with a scaffold size longer than 1 kbp, the number of false positives decreased by 93% (from 743 to 52) while the number of true positives decreased by only 17% (from 292 to 241) (Figure 2-b). This simulation clearly demonstrate that INTERCHANGE is able to efficiently identify simulated HTs directly from the raw reads without any detection bias towards genes or TEs. Additionally, the sensitivity of INTERCHANGE increases as the evolutionary distance between the species involved in HTs increases.

INTERCHANGE validation using real data allows detection of unknown horizontal transfer events

To further validate INTERCHANGE pipeline, we applied it to five distant plant genomes for which several HT have been previously reported [9]: grapevine (*Vitis vinifera*), peach (*Prunus persica*), poplar (*Populus trichocarpa*), date palm (*Phoenix dactylifera*) and clementine (*Citrus clementina*). These highly divergent species have experienced 6 HTs of LTR-

retrotransposons named BO1, BO2, BO3, BO4, BO7 and BC1 (BO : HT between plant orders; BC : HT between plant classes) [9]. In this previous study, the identification of these HTs was done through a comparative genomic analysis using assembled and previously annotated genomes. Here, we used the unassembled short reads of the same species (Supplemental Table 1) to test whether INTERCHNAGE could detect the previously reported HTs using a minimum scaffold size filter of 1 kbp to reduce the number of false positives as suggested by our HT simulation.

A total of 10 whole genome comparisons were performed between the 5 species and 31 HTs candidates were identified using INTERCHANGE, of which 30 correspond to LTR retrotransposons (29 *Copia* and 1 *Gypsy*) and one single gene (*Elongation factor 1*) (Supplemental Table 2). In addition to the HS criteria used by INTERCHANGE, we tested for the PI and PD criteria (see Method). The PD criterion was only tested if the HS and PI criteria are satisfied. Nine HT candidates meet both the HS, PI and PD criteria, while for the other 22, only one criterion was met leading to their rejection for further analysis (Supplemental Table 2). Four among the six known HTs (BO1, BO3, BO4 and BO6) were identified by our new strategy (see Figure 3). BO2 and BC1 were not detected by INTERCHANGE because they did not pass the 1kbp scaffold size filter (Supplemental Table 2). Remarkably, INTERCHANGE detect an additional five HTs that were not previously identified by El Baidouri et al. (2014). This include two HTs between grapevine and date palm (HT1, HT2), one HT between grapevine and poplar (HT3) and two HTs between poplar and prunus (HT4, HT5). Interestingly, as with the previously identified HTs, these new HTs correspond to LTRs-retrotransposons from the *Copia* superfamily. Strikingly, the LTR-retrotransposons identified by INTERCHANGE as having been transferred between grapevine and date palm (HT1 and HT2) have a high degree of sequence identity at 91% and 95.5%, respectively, suggesting a more recent transfer than the previously identified transfer (86% for BC1) (Supplemental Table 2). All newly identified HTs were also found in the reference genomes of species involved in HTs, providing further evidence of the reliability of the INTERCHANGE pipeline.

Characterization of horizontal transfers between wild non-model species from a natural ecosystem

In order to better understand the nature of biotic interactions that can promote HTs in natural ecosystems and whether some particular genes or TEs are more prone to HT than others, we

chose the Massane forest, a protected reserve in southern France and a unique site in Europe designated as a UNESCO World Heritage Site in July 2021. It is mainly composed of beech trees (*Fagus sylvatica*) located between 600 and 1158 m of altitude and covering 336 hectares in total. We sampled 17 different species from this ecosystem including 14 plant species and 3 fungi (see Table 1). The plant species selected include 6 tree species, 4 lianas/climbing plants, 2 herbaceous, 1 bramble, and 1 shrub and are all non-parasitic species. The 3 fungi species were collected from tree trunks or dead wood and include 2 sparophitic and 1 parasitic species. The selection of these species was based on a combination of biological and technical criteria, including: (i) the nature of their biotic interactions, including those with close relationships such as parasitism or physical proximity, and those with no known close interactions; (ii) genome size, with the aim of obtaining sufficient sequencing coverage for the detection of HTs. The selected species have a genome size smaller than 3 Gbp; (iii) abundance in the Massane forest, to facilitate collection; and (iv) taxonomic diversity to optimize phylogenetic representation.

The genomes of the selected species were sequenced using *Illumina* short-reads technology with 20X coverage (see Method and Supplemental Table 3). Using INTERCHANGE, we performed 136 whole-genome pairwise comparisons to identify highly similar regions between these species that may have originated from HTs. INTERCHANGE detected 68 HT candidates comprising 46 TEs and 22 genes and involving 8 out of the 17 studied species (see Supplemental Table 4). In order to avoid redundancy of candidates due to the presence of multiple paralogs, clustering was performed using SiLiX [32], resulting in 48 HT clusters (see Supplemental Table 4). To test the PI criteria, HT candidates where aligned to 400 plant genomes using Blastn (See Supplemental Table 5). Phylogenetic trees of the transferred TEs were constructed and compared to the phylogenetic trees of species (see Method Step 9). In total, of the 48 HT candidates (22 genes and 25 TEs) that met the HS criteria only 12 TE candidates also met the PI criteria (Supplemental Figure 2 to 13). 11 of the 12 TE candidates belongs to the LTR Copia superfamily (named MaCo01 to MaCo12, for Massane Copia) and 1 element belongs to the Gypsy superfamily (MaGy01, for Massane Gypsy). We further checked the presence/absence of these TEs in the genome 400 plant genomes to test for the PD criteria. This analysis clearly showed that these elements have a patchy distribution, thus confirming the occurrence of HT (see Supplemental Figure 14, to 24). Notably, both the phylogenetic trees and the patchy distribution of these transferred TEs in 400 plant species

point to other possible HT of the same elements between multiple plant species, suggesting that these elements may have undergone multiple HTs during their evolution.

Wet-lab validation of the horizontally transferred LTR-retrotransposons

To rule out potential contamination between the investigated species, despite all the precautions taken during sampling (see Methods), we performed PCR and sanger sequencing to check the presence of the identified HTs in the genome of the 8 species involved in HTs. To do this, we re-sampled two additional individuals from each of the 8 species and extracted their DNA. For each transferred LTRs, a set of two primer pairs was designed to amplify different regions of each element (Supplemental Table 6, Supplemental Figure 26 and Supplemental Figure 27). In the two individuals of the species involved in the HTs, the transferred LTRs were successfully amplified. For one candidate (MaCo11), only one of the two primer sets results in PCR amplicon (Supplemental Figure 26). One PCR amplicon from each transferred LTR was selected for sanger sequencing. Multiple sequence alignment of the PCR product sequences and the sequences of the elements detected using INTERCHANGE, as well as the construction of a phylogenetic tree, validate the occurrences of these HTs (see Data Availability). We also verified the presence of the transferred LTRs in the genomes of the donor/recipient species for which a reference genome is available. These are *A. glutinosa*, *F. sylvatica*, *F. excelsior* and *P. avium*. All the transferred LTRs implicating one of these species where identified in their respective reference genome. Additionally, we sampled and sequenced the genome of two *H. helix* individuals (Ivy A : ∼4Gb ; N50=14.4 kbp and Ivy B : ∼4Gb ; N50=14.6 kbp) and one *F. sylvatica* (7.4 Gb ; N50=20 kbp) using Nanopore sequencing (see Method section). We were also able to unambiguously identify the transferred LTRs involving these two species in different Nanopore reads corresponding to different paralogs (see Supplemental Figure 28 and 29). Taken together, these results clearly refute the possibility that the HTs identified in this study are the result of contamination.

Species involved and age of horizontal transfers

Among the 12 HTs that we identified, none involved a saprophytic or parasitic fungi. As shown in Figure 3, these HTs occurred between 8 out of the 17 studied species. The species involved in these transfers are essentially trees and climbing plants. Indeed, 5 of the 6 analyzed tree species have experienced at least one HT event. These are, in decreasing order of HT frequency: *Fraxinus excelsior* (6 HTs), *Fagus sylvatica* (5 HTs), *Alnus glutinosa* (2 HTs), *Acer monspessulanum* (2 HTs), *Prunus avium* (1 HT). For the climbers, two species

among the five analyzed have undergone HTs, namely *Dioscorea communis* (5 HTs) and *Hedera helix* (2 HTs). These HTs were identified between phylogenetically distant species that do not belong to the same plant class. In particular, the five HTs involving *D. communis* (Figure 4), which is a monocot species, occurred with eudicot species that diverged over 150 million years ago. Interestingly, most HTs involving *D. communis* (4/5 HTs) and *H. helix* (2/2 HTs) occurred with tree species which may suggest that the close physical relationship between lianas and trees may be a facilitator of HTs between those plants. Additionally, we found that some species pairs underwent multiple independent HTs of different LTR families such as the ones that occurred between *D. communis* and *F. excelsior* (2 HTs) and between *F. excelsior and F. sylvatica* (2 HTs). The direction of the HTs could not be determined, although the patchy distribution is clearly shown for all transferred LTRs (Supplemental Figures 14-25), as this will require further sampling and sequencing of additional plant genomes.

Sequence identity between the transferred LTRs varies from 89 to 97% (Supplemental Table 4), corresponding to an age of transfer between 3.8 and 1.15 million years (Mya) (using the molecular clock rate of Ma and Bennetzen, 2004)[33]. This indicates that these HTs occurred millions years ago and are therefore ancients. This is also supported by the PCR analysis on different individuals that indicate that these HTs are likely to be fixed in populations of these species. Additionally as shown earlier, for the species for which the reference genome is available (*F. sylvatica, F. excelsior, P. avium* and *A. glutinosa*) the transferred elements were found in their respective genomes pointing to ancient HT events. From the available data, it cannot be determined where these HTs took place and whether they occurred in the Massane forest, despite its ancient origin. It is important to note however that all species involved in these HTs are native to European and Mediterranean regions and that their respective geographic distributions overlaps, indicating that they have been in contact for a long period of time, thus facilitating the occurrence of HTs.

Copia LTRs-retrotransposons are the most frequently horizontally transferred elements in the investigated plant species

Despite the fact that our approach of HT identification does not focus on specific types of sequences such as TEs or genes, unlike all other approaches, the HTs identified in the Massane forest involve only LTRs-retrotransposons. Some might argue that INTERCHANGE has a bias toward LTR-retrotransposons transfer detection, but our simulations have shown

that this is not the case. Further characterization of the protein-coding genes of these transferred LTR-retrotransposons shows that 11 out of 12 belong to the Copia lineages (i.e. MaCo01 to MaCo11) and one belongs to the Gypsy lineage (MaGy01). This result is consistent with what we observed in our previous work, where *Copia* were more frequently transferred than *Gypsy* (28 *Copia* vs 7 *Gypsy*) [9]. However, in order to ascertain this, it is essential to check whether this is not due to an overrepresentation of the *Copia* superfamily among other LTRs in the surveyed genomes. For this purpose, we estimated the relative frequencies of *Copia* and *Gypsy* in the 8 species involved in HTs by aligning their raw genomic reads to a collection of reference protein sequences [34] using Diamond Blastx [35] (See Method). As shown in Figure 5-a, *Copia* elements were more prevalent than *Gypsy* elements in 6 out of the 8 species, equally abundant in *D. communis* and less prevalent in *H. helix*. On average, *Copia* were 1.4 times more prevalent than *Gypsy*. However this can not explain that 11 out of the 12 identified HTs belong to the *Copia* clade.

The transferred *Copia* LTR-retrotransposons belongs to Ale and Ivana lineage

We then investigated whether some *Copia* lineages have a greater propensity to transfer than others. To this end, we extracted from the Rexdb database the reverse transcriptase (RT) protein sequences of 17 *Copia* reference clades described in the literature [34] as well as those of the 11 transferred *Copia* (MaCo01 to MaCo11) identified in this study (21/22 paralogs) (see Method). Furthermore, we also extracted the RT domain of the 28 *Copia* previously identified as horizontally transferred between several plant species [9]. The constructed phylogenetic tree shows that for the previously described HTs, 21 *Copia* (75%) belong to only two lineages Ale (13) and Ivana (8) where the 7 remaining HTs belong to different lineages such as Tork, TAR or Bianca (Figure 5-b). This trend was even more pronounced for the *Copia* elements identified in the Massane forest. In fact, all transferred elements belong only to these two lineages: Ivana (6 /11) and Ale (5/11) (Figure 5-b).

These results suggest that these two lineages are more prone to HTs compared to other *Copia* lineages in the studied species. In order to check whether this observation could be due to an overrepresentation of these two *Copia* lineages in the analyzed plant genomes, we estimated the frequency of all known *Copia* clades in the 8 plant species involved in the identified HTs (see Method). This analysis shows big disparities in *Copia* lineages frequencies in those species with no particular conserved trend. In five of the eight species involved in HTs (*A. glutinosa*, *D. communis*, *F. sylvatica*, *P. avium*, *R. ulmifolius*) the Ale lineage was the

predominant *Copia* lineage ranging from 8.6% in *A. monspessulanum* to 37.8% in *P. avium* (Figure 4-b). On average Ale accounted for 24% of the *Copia* elements in these genomes, followed by SIRE lineage (16.6%). Meanwhile, the percentage of the Ivana lineage varies from 3.3% in *A. monspessulanum* to 17% in *D. communis* with an average of around 6.5%. These results show that the strong bias observed in transferred elements belonging to Ale and Ivana lineages cannot be explained by the relative abundance of these lineages in the genomes of the species involved in the HTs.

The horizontally transferred *Copia* are active after their transfer but show low transpositional activity in both donor and receiver species

To better understand the dynamics of the transferred elements, we estimated their copy number in both donor and receiver species using unassembled raw genome sequencing reads. To this end, we used the coverage of single-copy BUSCO genes as a standard to normalize the observed coverage of each transferred element (see Methods). This analysis show that copy number of the transferred elements varies from single copy to 28 copies for the Gypsy MaGy01 with an average of 4.3 copy per species (see Supplemental Table 7). For the HTs where we could identify the direction of HT, we can notice that some *Copia* did not transpose after their transfer (3/8) since they are single copies in the recipient species. The remaining elements show, on the contrary, a transpositional activity in the host recipient genomes that results in several copies. However, this post-transfer transpositional activity appears to be low, with only 2 to 5 copies observed for each transferred LTRs. It is interesting to note that this low transpositional activity is also observed in the donor species, suggesting that it is an inherent feature of the transferred *Copia element*.

Discussion

INTERCHANGE a new pipeline for HT characterization at the whole-genome scale using raw sequencing reads

In this study we investigated for the first time HTs between wild non-model species within a natural ecosystem. We sequenced the whole genome of 17 species including trees, climbing plants and fungi and characterized HTs directly form raw sequencing reads thanks to INTERCHANGE pipeline. Using this tool, we were able to report new HT events in previously studied species that had not been identified using previous methods [9]. In this study, we demonstrated the utility of INTERCHANGE for genome-wide screening of HT

events between non-model species, using both simulated and real datasets. Importantly, our approach obviates the need for costly and time-consuming genome assembly and annotation, which are often major bottlenecks for conducting such studies.

As indicated by our evaluation of INTERCHANGE some of the known or simulated HT even did not pass the 1 kbp scaffold size filter. A decrease in scaffold size may allow characterization of these events, but this will be at the expense of specificity. Other parameters may also impact the specificity or sensitivity of INTERCHANGE. For example, the smaller the k-mer size, the greater the sensitivity and vice versa, but this will lead to an increase in the computational time needed to perform all the possible pairwise comparisons. The different parameters of INTERCHANGE can be modified by the user which allows great flexibility. However, we shall point out that INTERCHANGE can only detect relatively recent HTs because high sequence divergence between older transferred sequences will not satisfy the HS criteria. Despite these limitations, our results show that this pipeline is very efficient at detecting gene and TEs HT events at the whole-genome scale using unassembled sequencing reads and is therefore a tool of choice for future studies of HT in natural ecosystems. INTERCHANGE could also be used to identify conserved sequences such as homologous genes, TEs or other types of sequences from unassembled genomes, which could be very useful for comparative genomics studies.

No plant-fungus horizontal transfer was identified at the Massane forest

Using the INTERCHANGE pipeline, we were able to identify 12 HTs implicating 8 plant species. We did not identify any transfer involving fungi even though the three studied species are saprophytic or parasitic and known to proliferate on tree trunks or dead wood. It is broadly accepted that close relationships such as endosymbiosis or parasitism are favorable for HTs in eukaryotes [19,21,24,36–38]. The absence of plant-fungus HT in this study may suggest that such events are rare or too old to be detected [39].

Liana-tree interactions: a possible new route of horizontal transfer between non-parasitic plants?

Our results also show that the two climbing plants, the common ivy (*H. helix*) and black bryony (*D. communis*) have experienced several HTs events predominantly with trees. These findings are in agreement with our previous study that showed a higher frequency of HTs between grapevines and several tree species [9]. To date, no hypothesis has been put forward

to explain this higher frequency of HT in grapevines and whether it is due to an inherent genetic trait or to its particular ecological lifestyle. In fact, similarly to common ivy and black bryony, wild grapevine use trees as support for growth which could explain the high HT frequency observed in this species. A recent study on four different and closely related *Vitis* species seems to confirm this trend [40]. Using comparative genomics, the authors identified dozens of HTs between these four closely related vine species and mainly trees belonging to highly divergent taxa, although they did not highlight the greater frequency of HTs between grapevine and trees. In light of our findings, we hypothesize that liana-tree interactions may favor HTs between non-parasitic plants and could be considered as route by which HTs occur frequently in nature. For the three other climbing plant genomes analyzed in this study, we did not find any HT. Therefore, the question of whether some climbing plants are more prone to HTs than others remains unanswered at this point. This needs to be tested on a larger sampling of liana species. It is also possible that the tree-to-tree HTs that we identified between beech (*F. sylvatica*), ash (*F. excelsior*) or alder (*A. glutinosa*) could be mediated by other, yet not sequenced, climbing plant species.

Horizontal transfers in plants mainly involve low copy number LTR-retrotransposons belonging to the Ivana and Ale lineages of the *Copia* superfamily

Our study reveals that LTR-retrotransposons are the only genetic elements that experienced HTs in the studied plant species, which confirms earlier reports, but remains without mechanistic explanation. In fact, successful HT requires three key steps, namely the "excision" of genetic material in the form of DNA or RNA molecules from the donor genome, its transport to the recipient species and finally its integration into the target genome. Due to their transposition life style, LTR-retrotransposons are able to generate extrachromosomal double stranded DNA encapsidated in the VLP (Virus Like Particule) and accumulating in the cytoplasm of the cells [41,42]. They also have the ability to integrate into the host genome using the integrase (IN) [43,44]. LTR-retrotransposons may therefore be more likely to achieve successful HT, given their ability to generate double stranded DNA encapsidated in the VLP and because of their ability to integrate the host genome. Although both *Copia* and *Gypsy* elements can produce VLPs, *Copia* appears to be more prone to horizontal transfer than *Gypsy* elements. If *Copia* and *Gypsy* superfamilies differ mainly in the order of the IN and RT domains, there are some genomic and transpositional features specific to each of these superfamilies. For instance, *Copia* elements are abundant in gene-rich euchromatic regions

while Gypsy elements are mainly located in Heterochromatic and pericentromeric regions [45]. *Copia* are also generally activated in response to environmental stress as it has been shown for many plant species [46]. It is therefore possible that Copia elements, because of their presence in transcriptionally active regions of the genome and because of their responsiveness to environmental stresses could facilitate their HT.

Intriguingly, *Copia* elements that have been horizontally transferred between plant genomes belong mainly to the Ivana and Ale lineage and are low copy numbers. The reasons why Ale and Ivana clade appear to be more prone to HT compared to the other *Copia* lineages remain unknown, as there are no well-known common specific features of these two clades that clearly differentiate them from others [47]. A recent population genomics study in *Arabidopsis arenosa* showed that *Copia* elements, particularly Ale and Ivana respond to temperature and irradiance [48]. It is also interesting to note that the TEs shown to be currently active in *A. thaliana* namely *EVADE* [49] and *ONSEN* [50], also belong to the Ale and Ivana clade, respectively, and the latter is active in response to heat stress [50]. As for the transferred *Copia*, these two families also have a low copy number: two copies for *EVADE* and eight copies for *ONSEN*. When a TE family reaches high copy numbers, it tends to be silenced by the production of small interfering RNAs and the epigenetic machinery depositing DNA methylation [51]. The silenced TEs would not be candidate for HT. If this holds true, the question of the presence and survival of Gypsy families in eukaryotic genomes remains to be explained by other mechanisms. Considering our study and previous ones on HT in plants, the propensity of *Copia* elements and in particular Ale and Ivana lineages to transfer horizontally can not be explained. Further studies are needed to elucidate the reasons for the remarkable ability of low copy number *Copia* to transfer horizontally in plants.

Conclusions

In this work, we conducted a pilot study on HTs *in natura* in a forest ecosystem. For this purpose, we implemented a new comparative genomics pipeline able to identify HTs at the whole genome level directly from raw sequencing reads. We characterized 12 HTs that all correspond to *Copia* LTRs retrotransposons and particularly those belonging to the Ale and Ivana lineages. Our study also shows that some lianas species have experienced recurrent horizontal transfers with trees that constitute their growth support in nature. This work sheds light on a new route of HTs between non-parasitic plant species and the type of genetic elements most likely to be horizontally transferred in plants.

Materiel and methods

Sampling

The 17 species analyzed were sampled in the Massane Forest National Nature Reserve. After sampling the target tissues (leaf or sporophore), the samples were first washed with a detergent solution (Tween 80 at 0.1%) and then rinsed twice successively in a miliQ water solution. The samples were then dried with absorbent paper and stored in liquid nitrogen and then at -80°C once in the laboratory before DNA extraction.

Illumina genome sequencing

DNA from each sample was extracted using the CTAB2X [52] and the quality of the DNA was estimated by Nanodrop (Thermo Scientific) and Qubit (Invitrogen) quantification. DNA libraries and sequencing was outsourced to Novogene company using the a Hiseq 2000 and Novaseq 6000 platforms. Briefly, a total amount of 1µg DNA per sample was used as input material for the DNA libraries. Sequencing libraries were generated using NEBNext® DNA Library Prep Kit following manufacturer's recommendations and indices were added to each sample. The genomic DNA was randomly fragmented to a size of 350bp by shearing, then DNA fragments were end polished, A-tailed, and ligated with the NEBNext adapter for *Illumina* sequencing, and further PCR enriched by P5 and indexed P7 oligos. The PCR products were purified (AMPure XP system) and resulted libraries were analyzed for size distribution by Agilent 2100 Bioanalyzer and quantified using real-time PCR. Paired-end sequencing was performed using a coverage of 20X and a read length of 150bp for each sample.

Nanopore genome sequencing

High-molecular-weight genomic DNA was extracted from 0.41 g to 0.5 g of frozen leaf tissue according to [53] and the Oxford Nanopore Technologies protocol (February 2019). Briefly, after lysis of cell membranes with Carlson buffer, proteins were removed with chloroform. DNA was purified using Qiagen Genomic-tip 100 columns following the manufacturer's instructions. A selection of fragments > 10 Kb was performed using AMPure XP beads. DNA quantification was performed by Nanodrop (Thermo Scientific) and Qubit assays (Invitrogen) and the quality was assessed on a 0.8% agarose gel. We then followed the 1D genomic DNA protocol by ligation with the SQK-LSK109 kit to prepare the 3 libraries using 3 µg, 3.9 µg,

and 4.1 µg of DNA (beech, ivy A, and ivy B), respectively. We successively loaded 1.7 µg of library onto a Flowcell R9, 2.6 µg and 2.7 µg of libraries onto two Flowcell R10. We produced 7.4 Gb, and 2 times 4 Gb of fastq pass reads with N50s of 20 kb, 14.4 kb and 14.6 kb, respectively. Bascalling was performed using guppy in the high accuracy (hac) mode (https://nanoporetech.com/nanopore-sequencing-data-analysis).

PCR and Sanger sequencing

We utilized the Taq DNA Polymerase 2x Master Mix RED to carry out our PCR amplification. The mix was prepared for each sample by adding 7.5 µl of Taq 2x Master Mix, 6.4 µl of H2O, 0.6 µl of the oligo F and R mix (10µM each), and 0.5 µl of DNA at 10ng/µl. The PCR program we employed consisted of a lid temperature of 98°C, followed by 34 cycles of denaturation at 95°C for 30 seconds, annealing at 60°C for 30 seconds, and extension at 72°C for 2 minutes. This was followed by a final extension at 72°C for 5 minutes and a cooling step at 4°C for 10 minutes. Finally, the samples were electrophoresed on a 1% agarose gel containing TAE 0.5X for 25 minutes at 135 volts, along with a 1kb+ size marker, to visualize the PCR products. PCR products have been purified and sent to Eurofinsgenomics for sequencing using the LightRun Tube platform.

Detection of Horizontal transfer using INTERCHANGE pipeline

High similarity criteria (Step1 to 8 using INTERCHANGE automatic pipeline)

Step 1 - Identification of homologous reads derived from conserved regions using a k-mers approach: *k*-mers indexes (k = 30) were generated using Tallymer mkindex option [54]. with default parameters except for: -mersize 30; minocc 1. The search for identical *k*-mers between each species pair was performed using Tallymer search option with the following parameters: -output qseqnum qpos counts sequence.

Step 2 - Once identical k-mer have been identified between reads of two species, the overlapping k-mer are merged and the total similarity score is calculated for each pair of reads using the following formula: Read similarity = total length of identical non-overlapping k-mer / reads length. Reads with a similarity score greater than 50% are considered to originate from conserved homologous regions and are therefore kept for further analysis.

Step 3 - There are a significant number of identical k-mers that correspond to regions of simple repeats such as tandem repeats. Reads containing such repeats are removed using

Prinseq-lite tool [55] with the following parameters: out_format 1; -lc-method dust; -lc-thresholds 10.

Step 4 - The homologous reads that pass the similarity filter are then extracted and assembled separately for each species using SPAdes [56] with the paird-end and only_assemble options. This step will result in the assembly in each species of scaffolds corresponding to highly conserved regions potentially derived from HTs.

Step 5 - The assembled scaffolds are then aligned using both Diamond blastp [35] and BLASTn against several databases with a minimum e-value de 1e-5 et 1e-20 respectively: CDDdelta, Repbase, mitochondrial, chloroplast, and ribosomal (TIGR) gene database [57].First, sequences that align to mitochondrial, chloroplastic and ribosomal genes are excluded. Indeed, these genes are generally highly conserved between distant species and therefore often meet the criterion of high similarity. When a scaffold aligns to several target sequences from multiple databases, only target sequences with the highest alignment score are considered as being homologous. At the end of this step, each scaffold will be classified into one of these categories: genes, TEs, MCRs (mitochondrial, chloroplast or ribosomal genes)

Step 6 - Identification of homologous scaffolds: the objective of this step is to identify homologous scaffolds between each pair of compared species. For this purpose, a reciprocal Blastn is performed and homologous scaffolds are identified using the reciprocal best hit method (RBH).

Step 7 - In order to distinguish, among the set of conserved scaffolds identified by INTERCHANGE, those that could originate from HT, it is necessary to first test the criterion of high similarity (HS). This means that the similarity of the transferred sequences between the donor and recipient species must be significantly higher than that of orthologous genes. Before assessing this criterion, it is therefore important to identify and assemble the conserved orthologous genes in the investigated species from unassembled short reads.

Characterization of orthologous BUSCO genes from unassembled reads: (i) Since the studied species from the Massane forest did not have any available gene annotation, we have assembled and annotated their BUSCO genes. These genes were used to test the HS criteria and to build the species phylogenetic tree. As a first step, the BUSCO genes of 400 publicly available assembled plant genomes (Supplemental Table 5) were identified, resulting in a genomic database of ~169,000 BUSCO genes covering angiosperms, gymnosperms and basal

plant species (this database has been deposited on the following link http://gamay.univ-perp.fr/~moaine/Database/). The genomic reads of each sequenced species from the Massane forest were mapped against this BUSCO database by minimap2 [58]. using default parameters. The mapped reads were extracted, merged and assembled by SPAdes [56] using paired-end and -only_assembler options. The resulting scaffolds were then realigned by Blastn against the nucleic BUSCO database and assigned to their corresponding BUSCO genes.

Step 8 - Identification of high sequence similarity threshold based on the distribution of orthologous gene identities: In order to identify whether conserved scaffolds have higher sequence similarity compared to orthologous BUSCO genes, a high similarity threshold (HS) is determined based on the distribution of orthologous gene sequence identities according to the following formula: HS = Q3+(IQR/2); where Q3 is the third quartile, IQR is the interquartile range (Q3-Q1).

Step 9 - Phylogenetic incongruence criteria (PI)

Building the phylogenetic tree of the studied species: the phylogenetic tree of the studied species is built based on BUSCO genes previously identified in step 7. Multiple alignment of orthologous BUSCO genes of the studied species and the 400 plant genomes is performed using Mafft program [59]. The alignments are then cleaned with TrimAl [60] and the trees constructed with FastTree [61]. A consensus tree is then obtained using Astral [62] from the previously constructed trees.

Building the phylogenetic tree of the transferred sequence: To construct the phylogenetic tree of the transferred elements, we aligned each of these elements to the assembled genomes of 400 plant species using Blastn. Sequences with sequence identity greater than 80% and covering at least 60% of the element were considered homologous. We performed multiple alignments for each element and all its homologs using the Mafft program [59]. These alignments were then cleaned using trimAL [60] with the following parameters: -cons 30; -gt 0.5. Finally, phylogenetic trees for each transferred element were inferred with FastTree [61]. The resulting trees were then manually compared to the species trees to check for the presence or absence of phylogenetic incongruencies. The trees were visualized using the Iroki Phylogenetic Tree Viewer [63]. The PI criterion is met when the phylogenetic tree of the HT candidate shows that the donor and recipient species are sister clades, unlike the species tree.

Step 10- Testing the Patchy distribution (PD): Finally, to consider that there is an unequal distribution of this sequence in the tree of species, the candidate sequence must be found in species close to donor/recipient but missing in species closely related to partner implicated in the HT. Alternatively, the transferred sequence could also be found only in the two species involved in the transfer (Supplemental Figure 14 to 25).

Candidates meeting the HS, the PI and the PD criteria are therefore considered as resulting from HTs.

In silico simulation of horizontal transfer events

In this simulation, we randomly selected 100 genes and 100 Tes from each species and randomly introduced them into the reference genomes of the other species. For TEs, we selected different copies belonging to the major classes (Class I and Class II) and to various TEs superfamilies (LTRs, LINEs, MuDR, hAT, Mutator, Helitrons...etc) with an equivalent proportion when possible. Before inserting these genes and TEs into the recipient genome, we artificially introduced mutations to create sequence divergence, simulating both recent and ancient HTs. The mutated sequences had a sequence divergence ranging from 80% to 100% identity compared to the original copies in the donor species. Wgsim tool (https://github.com/lh3/wgsim) was used to simulate 150 bp length paired-end reads with 20X coverage from the donor and recipient genomes carrying the *in silico* HTs using default parameters.

Estimation of copy LTRs number in unassembled genomes

To estimate the copy number of each retrotransposon in the species involved in the transfer, we calculated the number of mapped reads on each transferred retrotransposon compared to the numbers of mapped reads on single-copy genes. Total reads for each species were mapped onto the transferred LTRs-retrotransposons. For each LTR-retrotransposon, we calculated the coverage at each nucleotide of the element. The median coverage was taken as a proxy for the coverage of the element in the genome. The same strategy was adopted to estimate the coverage of the BUSCO genes of the studied species. We then used the following formula to estimate the total copy number of each transferred LTRs using *Illumina* reads: Copy number = (MCT / MCB), where MCT is the median coverage of LTRs and MCB: median coverage of BUSCO genes. To test whether this approach is an appropriate method to estimate copy number using genomic raw reads, we compared the copy number estimated from unassembled

genomes with that obtained from assembled reference genomes in species for which the latter is available (Supplemental Table 5). Copy numbers estimated from unassembled genomes and those obtained by Blastn against reference genomes are highly correlated, validating our approach (Pearson correlation; R = 0.982, p-value = 4.699E-10).

Phylogenetic tree of copia lineages

We extract the RT (reverse transcriptase) domain of the transferred *Copia* elements in both donor and receiver species (22 paralogs corresponding to the 11 *Copia* families). For 60% of the paralogs (13/22), the RT domain was assembled using our automatic INTERCHANGE pipeline. For the others, the RT domain was lacking. We then manually reassemble the lacking RT domains using raw *Illumina* reads of the corresponding species. For species for which the reference genome is available (*F. sylvatica*, *F. exclesior*, *P. avium* and *A. glutinosa*), we realigned the raw reads to the reference genomes and used the reference elements as a guide for manual assembly. Alternatively, we used homologs from other closely related plant species to guide manual assembly. Using this strategy, we obtained for most *Copia* paralogs involved in HTs nearly the complete elements with the corresponding RT domain (21/22).

Frequency estimates of the different *Copia* and *Gypsy* lineages

To estimate the relative frequency of *Copia* and *Gypsy* in the sequenced genomes, we aligned the raw genome reads of each species to a collection of protein sequences corresponding to the different known *Copia* and *Gypsy* lineages from the RexDB [34] database by Diamond Blastx (evalue 1e-5) [35]. The number of aligned reads on each superfamily and on each lineage was reported to the total number of aligned reads to estimate their relative frequency.

Data Availability

The data have been deposited in NCBI under BioProject accession number PRJNA788424 in the NCBI BioProject database (https://www.ncbi.nlm.nih.gov/bioproject/PRJNA788424). INTERCHANGE is open source and available at https://github.com/emaubin/INTERCHANGE. Multifasta sequences of 12 transferred LTRS (MaCo1 to MaCo12 and MaGy01) are available on this link (http://gamay.univ-perp.fr/~moaine/MaCo-MaGy/). PCR product sequences and multiple alignment of each of the 12 HTs are available on the following link (http://gamay.univ-perp.fr/~moaine/PCR/).

Funding

M.E is supported by a grant from the Agence Nationale de la Recherche (ANR-21-CE02-0031-01) and by the CNRS "diversity of biological mechanisms" call. O.P is supported by a BQR grant of the University of Perpignan, the Institut Universitaire de France and Occitanie region (Biodivoc). E.A is supported by a PhD grant from Occitanie region. M.M is supported by a grant from the Agence Nationale de la Recherche (ANR-21-PRCI-CE02). This study is set within the framework of the "Laboratoire d'Excellence (LABEX)" TULIP (ANR-10-LABX-41) and of the "Ecole Universitaire de Recherche (EUR)" TULIP-GS (ANR-18-EURE-0019).

Conflict of interest

The authors declare no conflict of interest.

Author contributions

M.E and O.P conceived and designed the study. J.G, E.A, M.E and O.P collected the samples. E.A and M.E performed the analysis. E.A and C.L extracted DNA for *Illumina* sequencing. C.L and M.M performed Nanopore sequencing. M.E wrote the manuscript with input from E.A. All authors provided critical feedback on the manuscript.

Plant sampling statement

The authors declare that the samples taken in the framework of this study comply with local and national legislation and that an authorization has been granted.

Acknowledgements

We are grateful to Marie-Christine Carpentier, Joris Bertrand for their help for plant sampling at the Massane forest.

Figure legend and table

Common Name	Species	Type	Estimaed Genome size	Available reference genome
Beech	*Fagus sylvatica*	Tree	540 Mbp	yes
Ash	*Fraxinus excelsior*	Tree	840 Mbp	yes

Montpellier maple	*Acer monspessulanum*	Tree	730 Mbp	no
Wild cherry	*Prunus avium*	Tree	430 Mbp	yes
Alder	*Alnus glutinosa*	Tree	500 Mbp	yes
Whitebeam	*Sorbus aria*	Tree	1.03 Gbp	no
Hairy Greenweed	*Genista pilosa*	Shrub	1.04 Gbp	no
Honeysuckle	*Lonicera periclymenum*	Liana	2.81 Gbp	no
Ivy	*Hedera helix*	Liana	1.5 Gbp	no
Black bryony	*Dioscorea communis*	Liana	830 Mbp	no
Giant blackberry	*Rubus ulmifolius*	Bramble	450 Mbp	no
White Bryony	*Bryonia dioica*	Liana	1.6 Gbp	no
Sage	*Salvia sp*	Herbaceous	760 Mbp	-
Tinder Bracket	*Fomes fomentarius*	Fungi	50 Mbp	yes
Coral Tooth	*Hericium clathroides*	Fungi	40 Mbp	no
Oyster mushroom	*Pleurotus ostreatus*	Fungi	20/50Mbp	yes
Narrow-leaved Ragwort	*Senecio inaequidens*	Herbaceous	580 Mbp	no

Table 1: Species sampled in the Massane forest and whose genome has been sequenced by *Illumina* short read sequencing.

Figure 1: The different steps of the INTERCHANGE pipeline of horizontal transfer identification from unassembled and unannotated genomes. Steps 1 to 8 are completely automatic steps 9 and 10 are semi-automatic. Step1: Identification of identical k-mers using Tallymer [54]. Steps 2 & 3 Identification of reads derived from conserved regions & Removal of Tandem repeats using PRINSEQ lite [55]. Reads sharing at least 50% of identical k-mers are considered as homologous reads. Step 4: homologous reads are extracted and assembled for each pair of species using SPAdes [56]. Step 5: Scaffolds annotation using multiple protein and TEs database: CDDdelta, Repbase, mitochondrial, chloroplast, and ribosomal (TIGR) gene database. Step 6: Identification of homologous scaffolds using reciprocal best hit (RBH). Step 7: Identification of high sequence similarity threshold based on the distribution of orthologous BUSCO gene identities according to the following formula: high similarity threshold (HS) = Q3+(IQR/2); where Q3 is the third quartile, IQR is the interquartile range

(Q3-Q1). Step 8: Testing for HS criteria. Step 9: Phylogenetic incongruence criteria. Step 10: testing the Patchy distribution (PD) of transferred sequence. For details see Method section.

Figure 2: Simulation of horizontal transfer (HT) between *A. thaliana* and *O. sativa* and between *O. sativa* and B. distachyon. a. 200 HT events were simulated in each direction (green arrows), comprising genes and TEs with equal proportion. b. INTERCHANGE results using short reads of genomes harboring simulated HTs. Y-axis indicate to the total number of HTs (scaffolds) identified by INTERCHANGE and X-axis represent filters based on scaffold size. The color codes are provided in the figure legend.

Figure 3: HTs identified by INTERCHANGE using real data. Lines represent the HT events identified from genome short read sequencing data. In green, HTs that were identified in a previous study [9] using reference genome and detected by INTERCHANGE from short reads. In gray, HTs missed by INTERCHANGE. In red, new HTs only identified by INTERCHANGE.

Figure 4: The phylogenetic tree of the 17 analyzed Massane species. The curves represent the identified HTs and link the involved species. Blue and red curves represent *Gypsy* and *Copia* HTs, respectively. The asterisks indicate multiple HTs. The horizontal scale represents the divergence time in million years (source: timetree.org). Correspondence of species names: Ace: *Acer monspessulanum*, Aln: *Alnus glutinosa*, Bry: *Bryonia dioica*, Dio: *Dioscorea communis*, Fag: *Fagus sylvatica*, Fra: *Fraxinus excelsior*, Fom: *Fomes fomentarius*, Gen: *Genista pilosa*, Hed: *Hedera helix*, Her: *Hericium clathroides*, Lon: *Lonicera periclymenum*, Ple: *Pleurotus ostreatus*, Pru: *Prunus avium*, Rub: *Rubus ulmifolius*, Sal: *Salvia sp*, Sen: *Senecio inaequendis*, Sor: *Sorbus aria*.

Figure 5: The relative abundance of LTRs-retrotransposon superfamilies in species that have experienced HTs. a) Relative frequency of Copia and Gypsy in the studied species involved in HTs. In blue: *Copia* frequency, in yellow: *Gypsy* frequency b) Phylogenetic tree of transferred *Copia* detected in this study using the RT domain. In bold, the consensus sequence of the reference *Copia* lineages. Maco1 to 11 correspond to horizontally transferred elements identified between the plant species from the Massane. BO1 to BO8, BG1 to BG and BC1 correspond to *Copia* elements identified in our previous study. Correspondence of species names as in Fig. 3. c) Copia lineages relative frequencies in species involved in HT.

Supplemental Figure 1: Sequence identity distribution of the assembled BUSCO genes in the studied species. These sequence identities were obtained by Blastn alignment. N: corresponds to the total number of BUSCO genes that can be aligned at the nucleic level between each pair of species. HS threshold calculated by INTERCHANGE using the following formula : HS = (Q3+IQR/2), the inter-quartile range IQR = Q3-Q1 (Q1 and Q3 correspond to the first and third quartile respectively). The age in millions of years (Mya) represents the divergence time between species according to Timetree.org. Correspondence of species names: Ace: *Acer monspessulanum*, Aln: *Alnus glutinosa*, Bry: *Bryonia dioica*, Dio: *Dioscorea communis*, Fag: *Fagus sylvatica*, Fra: *Fraxinus excelsior*, Fom: *Fomes fomentarius*, Gen: *Genista pilosa*, Hed: *Hedera helix*, Her: *Hericium clathroides*, Lon: *Lonicera periclymenum*, Ple: *Pleurotus ostreatus*, Pru: Prunus avium, Rub: *Rubus ulmifolius*, Sal: *Salvia sp*, Sen: *Senecio inaequendis*, Sor: *Sorbus aria*.

Supplemental Figure 2-13: Phylogenetic tree of the horizontally transferred LTR-retrotransposons constructed using all homologous elements from 400 plant species. MaCo01 to MaCo11 correspond to the transferred Copia superfamily and MaGy01 to the Gypsy superfamily. Nodes supported with boostrap values above 70% are indicated with a black dot. Nodes with bootstrap values under 70% are indicated with white dot.

Supplemental Figure 14-25: Patchy distribution of the horizontally transferred LTR-retrotransposons in the phylogenetic tree of 400 plant species. MaCo01 to MaCo11 correspond to the transferred Copia superfamily and MaGy01 to the Gypsy superfamily. The green bars represent the species harboring the LTR family and their height the relative abundance in the host genome.

Supplemental Figure 26-27: PCR validation of the transferred LTRs MaCo01 to MaCo11 and MaGy01. The red arrows indicate the primers designed to amplify different regions of the transferred LTRs in the species involved in the HTs. For each species, PCR was performed using DNA from two different individuals, different from those used for genome sequencing, to limit possible contamination.

Supplemental Figure 28 : Graphical visualisation of Blastn alignment of transferred LTRs-retrotransposons (Maco2) identified and assembled using INTERCHANGE pipeline against Nanopore reads of two *Hedera helix* genomes corresponding to two ivy individuals A and B. Visual representation was achieved using http://kablammo.wasmuthlab.org/ software.

Supplemental Figure 29: Graphical visualisation of Blastn alignment of two transferred LTRs-retrotransposons (Maco3 and Maco11) identified and assembled using INTERCHANGE pipeline against Nanopore reads of *Fagus sylvatica* genome. Visual representation was achieved using http://kablammo.wasmuthlab.org/ software.

Supplemental Table 1 : List of five plant species used as control data and their corresponding sequence read archive IDs used as input in INTERCHANGE.

Supplemental Table 2 : List of HT candidates detected by INTERCHANGE between the five species listed in Supplemental Table 1.

Supplemental Table 3 : List of species sequenced in the frame of this study and their BioSample IDs.

Supplemental Table 4 : List of HT candidates detected (MaCo1 to MaCo11 and MaGy01) between the 17 studied species. In green candidates meeting the HS, PI and PD criteria. In gray candidates meeting only the HS criteria.

Supplemental Table 5 : List of 400 plant species in which homologs of the transferred LTRs were screened.

Supplemental Table 6 : Primers used for PCR validation of transferred LTRs. For each candidate, two sets of primers were designed and labeled P1 and P2 (see Supplemental Figure 26 and 27).

Supplemental Table 7 : List of the 12 transferred LTRs, their size and copy number estimated from the unassembled genomes and from the reference genome when available.

References

1. Wiedenbeck J, Cohan FM. Origins of bacterial diversity through horizontal genetic transfer and adaptation to new ecological niches. FEMS Microbiology Reviews. 2011;35:957–76.

2. Davies J. Inactivation of antibiotics and the dissemination of resistance genes. Science. 1994;264:375–82.

3. Keeling PJ, Palmer JD. Horizontal gene transfer in eukaryotic evolution. Nature Reviews Genetics. 2008;9:605–18.

4. Boto L. Horizontal gene transfer in the acquisition of novel traits by metazoans. Proceedings of the Royal Society B: Biological Sciences [Internet]. 2014;281:20132450. Available from: https://royalsocietypublishing.org/doi/10.1098/rspb.2013.2450

5. Soucy SM, Huang J, Gogarten JP. Horizontal gene transfer: Building the web of life. Nature Reviews Genetics [Internet]. Nature Publishing Group; 2015;16:472–82. Available from: http://dx.doi.org/10.1038/nrg3962

6. Sibbald SJ, Eme L, Archibald JM, Roger AJ. Lateral Gene Transfer Mechanisms and Pan-genomes in Eukaryotes. Trends in Parasitology [Internet]. 2020;36:927–41. Available from: https://linkinghub.elsevier.com/retrieve/pii/S1471492220301975

7. Van Etten J, Bhattacharya D. Horizontal Gene Transfer in Eukaryotes: Not if, but How Much? Trends in Genetics [Internet]. 2020; Available from: https://linkinghub.elsevier.com/retrieve/pii/S0168952520302067

8. Chen R, Huangfu L, Lu Y, Fang H, Xu Y, Li P, et al. Adaptive innovation of green plants by horizontal gene transfer. Biotechnology Advances [Internet]. 2021;46:107671. Available from: https://linkinghub.elsevier.com/retrieve/pii/S0734975020301737

9. El Baidouri M, Carpentier M-C, Cooke R, Gao D, Lasserre E, Llauro C, et al. Widespread and frequent horizontal transfers of transposable elements in plants. Genome Research [Internet]. 2014;24:831–8. Available from: http://genome.cshlp.org/lookup/doi/10.1101/gr.164400.113

10. Peccoud J, Loiseau V, Cordaux R, Gilbert C. Massive horizontal transfer of transposable elements in insects. Proceedings of the National Academy of Sciences [Internet]. 2017;114:4721–6. Available from: http://www.pnas.org/lookup/doi/10.1073/pnas.1621178114

11. Dunning LT, Olofsson JK, Parisod C, Choudhury RR, Moreno-Villena JJ, Yang Y, et al. Lateral transfers of large DNA fragments spread functional genes among grasses. Proceedings of the National Academy of Sciences [Internet]. 2019;201810031. Available from: http://www.pnas.org/lookup/doi/10.1073/pnas.1810031116

12. Zhang H-H, Peccoud J, Xu M-R-X, Zhang X-G, Gilbert C. Horizontal transfer and evolution of transposable elements in vertebrates. Nature Communications [Internet]. 2020;11:1362. Available from: http://www.nature.com/articles/s41467-020-15149-4

13. Christin PA, Edwards EJ, Besnard G, Boxall SF, Gregory R, Kellogg EA, et al. Adaptive evolution of C 4 photosynthesis through recurrent lateral gene transfer. Current Biology [Internet]. Elsevier Ltd; 2012;22:445–9. Available from: http://dx.doi.org/10.1016/j.cub.2012.01.054

14. Graham LA, Li J, Davidson WS, Davies PL. Smelt was the likely beneficiary of an antifreeze gene laterally transferred between fishes. BMC Evolutionary Biology [Internet]. BMC Evolutionary Biology; 2012;12:190. Available from: BMC Evolutionary Biology

15. Yue J, Hu X, Sun H, Yang Y, Huang J. Widespread impact of horizontal gene transfer on plant colonization of land. Nature Communications [Internet]. Nature Publishing Group; 2012;3:1152. Available from: http://dx.doi.org/10.1038/ncomms2148

16. Li FW, Villarreal JC, Kelly S, Rothfels CJ, Melkonian M, Frangedakis E, et al. Horizontal transfer of an adaptive chimeric photoreceptor from bryophytes to ferns. Proceedings of the National Academy of Sciences of the United States of America. 2014;111:6672–7.

17. Wang H, Sun S, Ge W, Zhao L, Hou B, Wang K, et al. Horizontal gene transfer of Fhb7 from fungus underlies Fusarium head blight resistance in wheat. Science [Internet]. 2020;eaba5435. Available from: https://www.sciencemag.org/lookup/doi/10.1126/science.aba5435

18. Xia J, Guo Z, Yang Z, Han H, Wang S, Xu H, et al. Whitefly hijacks a plant detoxification gene that neutralizes plant toxins. Cell [Internet]. Elsevier Inc.; 2021;1–13. Available from: https://doi.org/10.1016/j.cell.2021.02.014

19. Yang Z, Zhang Y, Wafula EK, Honaas LA, Ralph PE, Jones S, et al. Horizontal gene transfer is more frequent with increased heterotrophy and contributes to parasite adaptation. Proceedings of the National Academy of Sciences of the United States of America. 2016;113:E7010–9.

20. Yoshida S, Maruyama S, Nozaki H, Shirasu K. Horizontal gene transfer by the parasitic plant striga hermonthica. Science. 2010;328:1128.

21. Yang Z, Wafula EK, Kim G, Shahid S, McNeal JR, Ralph PE, et al. Convergent horizontal gene transfer and cross-talk of mobile nucleic acids in parasitic plants. Nature Plants [Internet]. Springer US; 2019; Available from: http://www.nature.com/articles/s41477-019-0458-0

22. Davis CC. Host-to-Parasite Gene Transfer in Flowering Plants: Phylogenetic Evidence from Malpighiales. Science [Internet]. 2004;305:676–8. Available from: https://www.sciencemag.org/lookup/doi/10.1126/science.1100671

23. Xi Z, Bradley RK, Wurdack KJ, Wong KM, Sugumaran M, Bomblies K, et al. Horizontal transfer of expressed genes in a parasitic flowering plant. BMC Genomics [Internet]. 2012;13:227. Available from: http://bmcgenomics.biomedcentral.com/articles/10.1186/1471-2164-13-227

24. Gilbert C, Schaack S, Pace II JK, Brindley PJ, Feschotte C. A role for host–parasite interactions in the horizontal transfer of transposons across phyla. Nature [Internet]. 2010;464:1347–50. Available from: http://www.nature.com/articles/nature08939

25. Stegemann S, Bock R. Exchange of Genetic Material. Science. 2009;649.

26. Fuentes I, Stegemann S, Golczyk H, Karcher D, Bock R. Horizontal genome transfer as an asexual path to the formation of new species. Nature. 2014;511:232–5.

27. Hibdige SGS, Raimondeau P, Christin P-A, Dunning LT. Widespread lateral gene transfer among grasses. New Phytologist [Internet]. 2021;230:2474–86. Available from: https://onlinelibrary.wiley.com/doi/10.1111/nph.17328

28. Magri D. Patterns of post-glacial spread and the extent of glacial refugia of European beech (Fagus sylvatica). Journal of Biogeography [Internet]. 2008;35:450–63. Available from: http://doi.wiley.com/10.1111/j.1365-2699.2007.01803.x

29. Aubin E, El Baidouri M, Panaud O. Horizontal Gene Transfers in Plants. Life [Internet]. 2021;11:857. Available from: https://www.mdpi.com/2075-1729/11/8/857

30. Wallau GL, Ortiz MF, Loreto ELS. Horizontal Transposon Transfer in Eukarya: Detection, Bias, and Perspectives. Genome Biology and Evolution [Internet]. 2012;4:801–11. Available from: https://academic.oup.com/gbe/article-lookup/doi/10.1093/gbe/evs055

31. Ellinghaus D, Kurtz S, Willhoeft U. LTRharvest, an efficient and flexible software for de novo detection of LTR retrotransposons. BMC Bioinformatics. 2008;9.

32. Miele V, Penel S, Duret L. Ultra-fast sequence clustering from similarity networks with SiLiX. BMC Bioinformatics [Internet]. 2011;12:116. Available from: https://bmcbioinformatics.biomedcentral.com/articles/10.1186/1471-2105-12-116

33. Ma J, Bennetzen JL. Rapid recent growth and divergence of rice nuclear genomes. Proceedings of the National Academy of Sciences [Internet]. 2004;101:12404–10. Available from: http://www.pnas.org/cgi/doi/10.1073/pnas.0403715101

34. Neumann P, Novák P, Hoštáková N, Macas J. Systematic survey of plant LTR-retrotransposons elucidates phylogenetic relationships of their polyprotein domains and provides a reference for element classification. Mobile DNA [Internet]. 2019;10:1. Available from: https://mobilednajournal.biomedcentral.com/articles/10.1186/s13100-018-0144-1

35. Buchfink B, Xie C, Huson DH. Fast and sensitive protein alignment using DIAMOND. Nature Methods [Internet]. Nature Publishing Group; 2015;12:59–60. Available from: http://dx.doi.org/10.1038/nmeth.3176

36. Xi Z, Wang Y, Bradley RK, Sugumaran M, Marx CJ, Rest JS, et al. Massive Mitochondrial Gene Transfer in a Parasitic Flowering Plant Clade. PLoS Genetics. 2013;9:1–10.

37. Kim G, LeBlanc ML, Wafula EK, DePamphilis CW, Westwood JH. Genomic-scale exchange of

mRNA between a parasitic plant and its hosts. Science [Internet]. 2014;345:808–11. Available from: https://www.sciencemag.org/lookup/doi/10.1126/science.1253122

38. Leclercq S, Thézé J, Chebbi MA, Giraud I, Moumen B, Ernenwein L, et al. Birth of a W sex chromosome by horizontal transfer of Wolbachia bacterial symbiont genome. Proceedings of the National Academy of Sciences of the United States of America. 2016;113:15036–41.

39. Richards TA, Soanes DM, Foster PG, Leonard G, Thornton CR, Talbot NJ. Phylogenomic Analysis Demonstrates a Pattern of Rare and Ancient Horizontal Gene Transfer between Plants and Fungi. The Plant Cell [Internet]. 2009;21:1897–911. Available from: https://academic.oup.com/plcell/article/21/7/1897/6095451

40. Park M, Sarkhosh A, Tsolova V, El-Sharkawy I. Horizontal Transfer of LTR Retrotransposons Contributes to the Genome Diversity of Vitis. International Journal of Molecular Sciences [Internet]. 2021;22:10446. Available from: https://www.mdpi.com/1422-0067/22/19/10446

41. Lanciano S, Carpentier MC, Llauro C, Jobet E, Robakowska-Hyzorek D, Lasserre E, et al. Sequencing the extrachromosomal circular mobilome reveals retrotransposon activity in plants. PLoS Genetics. 2017;13:1–20.

42. Lee SC, Ernst E, Berube B, Borges F, Parent J-S, Ledon P, et al. Arabidopsis LTR retrotransposons and their regulation by epigenetically activated small RNA. 2020;576–88.

43. Sabot F, Schulman AH. Parasitism and the retrotransposon life cycle in plants : a hitchhiker ' s guide to the genome. 2006;381–8.

44. Lee SC, Martienssen RA. Regulation of retrotransposition in Arabidopsis. Biochemical Society Transactions [Internet]. 2021;49:2241–51. Available from: https://portlandpress.com/biochemsoctrans/article/49/5/2241/229743/Regulation-of-retrotransposition-in-Arabidopsis

45. Baduel P, Leduque B, Ignace A, Gy I, Gil J, Loudet O, et al. Genetic and environmental modulation of transposition shapes the evolutionary potential of Arabidopsis thaliana. Genome Biology [Internet]. 2021;22:138. Available from: https://genomebiology.biomedcentral.com/articles/10.1186/s13059-021-02348-5

46. Grandbastien M-A. LTR retrotransposons, handy hitchhikers of plant regulation and stress response. Biochimica et Biophysica Acta (BBA) - Gene Regulatory Mechanisms [Internet]. 2015;1849:403–16. Available from: https://linkinghub.elsevier.com/retrieve/pii/S1874939914002053

47. Wicker T, Keller B. Genome-wide comparative analysis of copia retrotransposons in Triticeae, rice, and Arabidopsis reveals conserved ancient evolutionary lineages and distinct dynamics of individual copia families. Genome Research [Internet]. 2007;17:1072–81. Available from: http://genome.cshlp.org/lookup/doi/10.1101/gr.6214107

48. Wos G, Choudhury RR, Kolář F, Parisod C. Transcriptional activity of transposable elements along an elevational gradient in Arabidopsis arenosa. Mobile DNA [Internet]. 2021;12:7. Available from: https://mobilednajournal.biomedcentral.com/articles/10.1186/s13100-021-00236-0

49. Mirouze M, Reinders J, Bucher E, Nishimura T, Schneeberger K, Ossowski S, et al. Selective epigenetic control of retrotransposition in Arabidopsis. Nature [Internet]. 2009;461:427–30. Available from: http://www.nature.com/articles/nature08328

50. Ito H, Gaubert H, Bucher E, Mirouze M, Vaillant I, Paszkowski J. An siRNA pathway prevents transgenerational retrotransposition in plants subjected to stress. Nature [Internet]. 2011;472:115–9. Available from: http://www.nature.com/articles/nature09861

51. Marí-Ordóñez A, Marchais A, Etcheverry M, Martin A, Colot V, Voinnet O. Reconstructing de novo silencing of an active plant retrotransposon. Nature Genetics [Internet]. 2013;45:1029–39.

Available from: http://www.nature.com/articles/ng.2703

52. Debladis E, Llauro C, Carpentier M-C, Mirouze M, Panaud O. Detection of active transposable elements in Arabidopsis thaliana using Oxford Nanopore Sequencing technology. BMC Genomics [Internet]. 2017;18:537. Available from: http://bmcgenomics.biomedcentral.com/articles/10.1186/s12864-017-3753-z

53. Pushkova E., Beniaminov A., Krasnov G., Novakovskiy R., Povkhova N., Melnikova N., et al. Extraction of high-molecular-weight DNA from poplar plants for Nanopore sequencing. Plant genetic resources for breeding and producing functional nutraceutical food [Internet]. Institute of Cytology and Genetics, Siberian Branch of the Russian Academy of Sciences Novosibirsk State University; 2019. p. 158–60. Available from: http://publ.icgbio.ru/wp-content/uploads/2019/10/158-160.pdf

54. Kurtz S, Narechania A, Stein JC, Ware D. A new method to compute K-mer frequencies and its application to annotate large repetitive plant genomes. BMC genomics. 2008;9:517.

55. Schmieder R, Edwards R. Quality control and preprocessing of metagenomic datasets. Bioinformatics [Internet]. 2011;27:863–4. Available from: https://academic.oup.com/bioinformatics/article-lookup/doi/10.1093/bioinformatics/btr026

56. Prjibelski A, Antipov D, Meleshko D, Lapidus A, Korobeynikov A. Using SPAdes De Novo Assembler. Current Protocols in Bioinformatics [Internet]. 2020;70. Available from: https://onlinelibrary.wiley.com/doi/abs/10.1002/cpbi.102

57. Ouyang S. The TIGR Plant Repeat Databases: a collective resource for the identification of repetitive sequences in plants. Nucleic Acids Research [Internet]. 2004;32:360D – 363. Available from: https://academic.oup.com/nar/article-lookup/doi/10.1093/nar/gkh099

58. Li H. Minimap2: pairwise alignment for nucleotide sequences. Birol I, editor. Bioinformatics [Internet]. 2018;34:3094–100. Available from: https://academic.oup.com/bioinformatics/article/34/18/3094/4994778

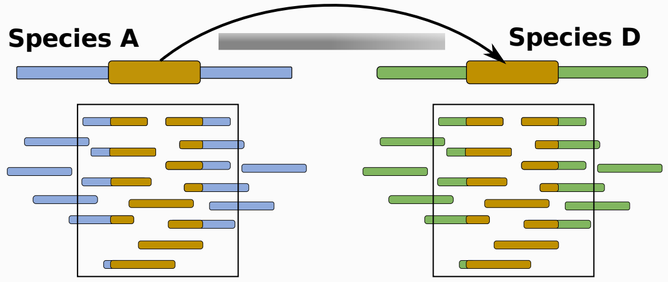
59. Katoh K, Standley DM. MAFFT Multiple Sequence Alignment Software Version 7: Improvements in Performance and Usability. Molecular Biology and Evolution [Internet]. 2013;30:772–80. Available from: https://academic.oup.com/mbe/article-lookup/doi/10.1093/molbev/mst010

60. Capella-Gutierrez S, Silla-Martinez JM, Gabaldon T. trimAl: a tool for automated alignment trimming in large-scale phylogenetic analyses. Bioinformatics [Internet]. 2009;25:1972–3. Available from: https://academic.oup.com/bioinformatics/article-lookup/doi/10.1093/bioinformatics/btp348

61. Price MN, Dehal PS, Arkin AP. FastTree: Computing Large Minimum Evolution Trees with Profiles instead of a Distance Matrix. Molecular Biology and Evolution [Internet]. 2009;26:1641–50. Available from: https://academic.oup.com/mbe/article-lookup/doi/10.1093/molbev/msp077

62. Zhang C, Rabiee M, Sayyari E, Mirarab S. ASTRAL-III: polynomial time species tree reconstruction from partially resolved gene trees. BMC Bioinformatics [Internet]. 2018;19:153. Available from: https://bmcbioinformatics.biomedcentral.com/articles/10.1186/s12859-018-2129-y

63. Moore RM, Harrison AO, McAllister SM, Polson SW, Wommack KE. Iroki: automatic customization and visualization of phylogenetic trees. PeerJ [Internet]. 2020;8:e8584. Available from: https://peerj.com/articles/8584



Step 1
Identification of identical *k*-mers (Tallymer)

Steps 2 & 3
Identification of reads derived from conserved regions & Removal of Tandem repeats (PRINSEQ lite)

Reads sharing identical *k*-mers over at least 50 % of read length

Step 4
Reads assembly (SPAdes)

Step 5
Scaffolds annotation (Blastn, Diamond blastp)

Step 6
Identification of homologous scaffolds

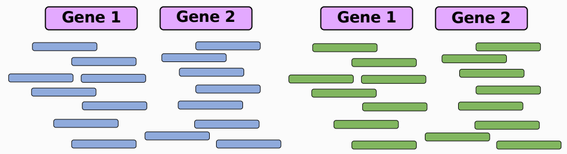
Reciprocal best hit

Alignement length> 1Kb

Step 7
Identification of BUSCO orthologous genes in non-assembled genomes

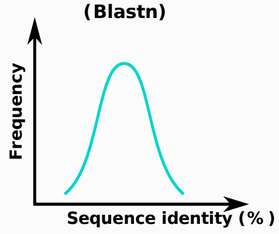
DB
BUSCO

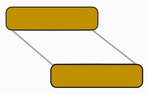
Alignment of reads against BUSCO nucleic database (Minimap2)



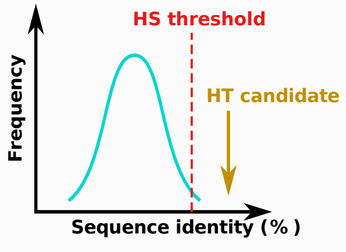
Assembly (SPAdes)

Orthologs gene pairs alignment (Blastn)

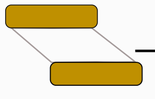




Step 8
HS test
Sequence identity of candidates > HS



HS validation



INTERCHANGE

horIzoNtal TransfER CHAracterization in Non-assembled GEnome

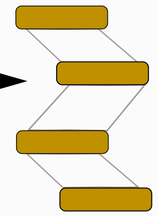
Step 9
Phylogenetic incongruence (PI)

DB
Plant Genomes (B, C, E, F, G, H, I ...etc)

Integration of all available plant genomes for PI test

Homologs identification in 400 plant genomes (BLASTn)

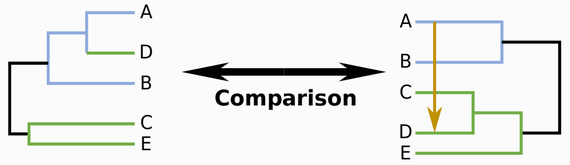
BUSCO genes from 400 plant genomes + BUSCO genes from studied species



Building phylogenetic tree of each ortholgous group (Fastree)

Phylogenetic tree of transfered sequence (Fastree)

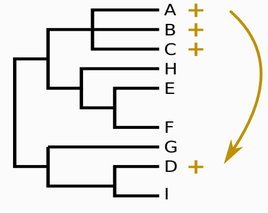
Consensus phylogenetic tree (ASTRAL)

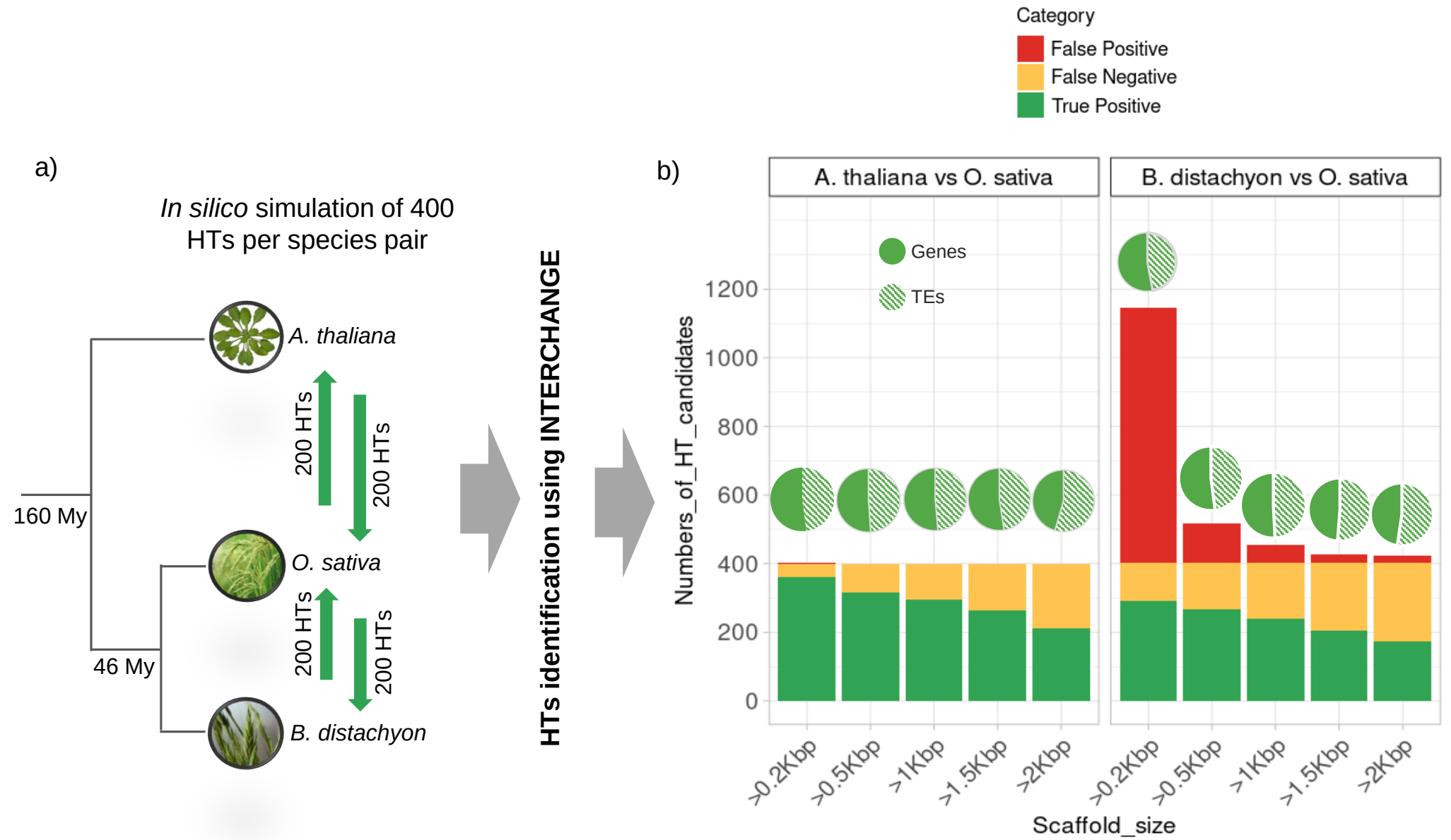


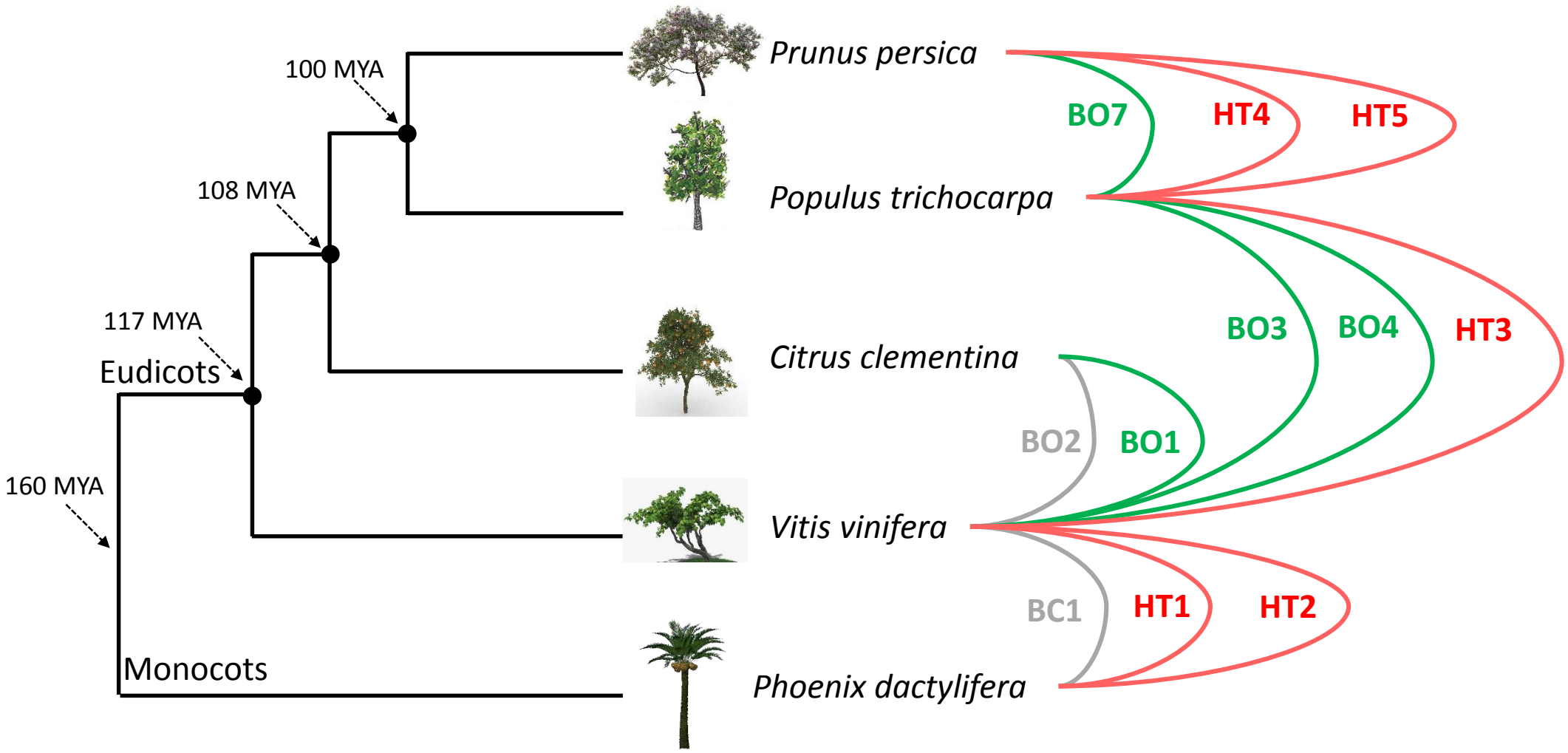
Transferred sequence tree

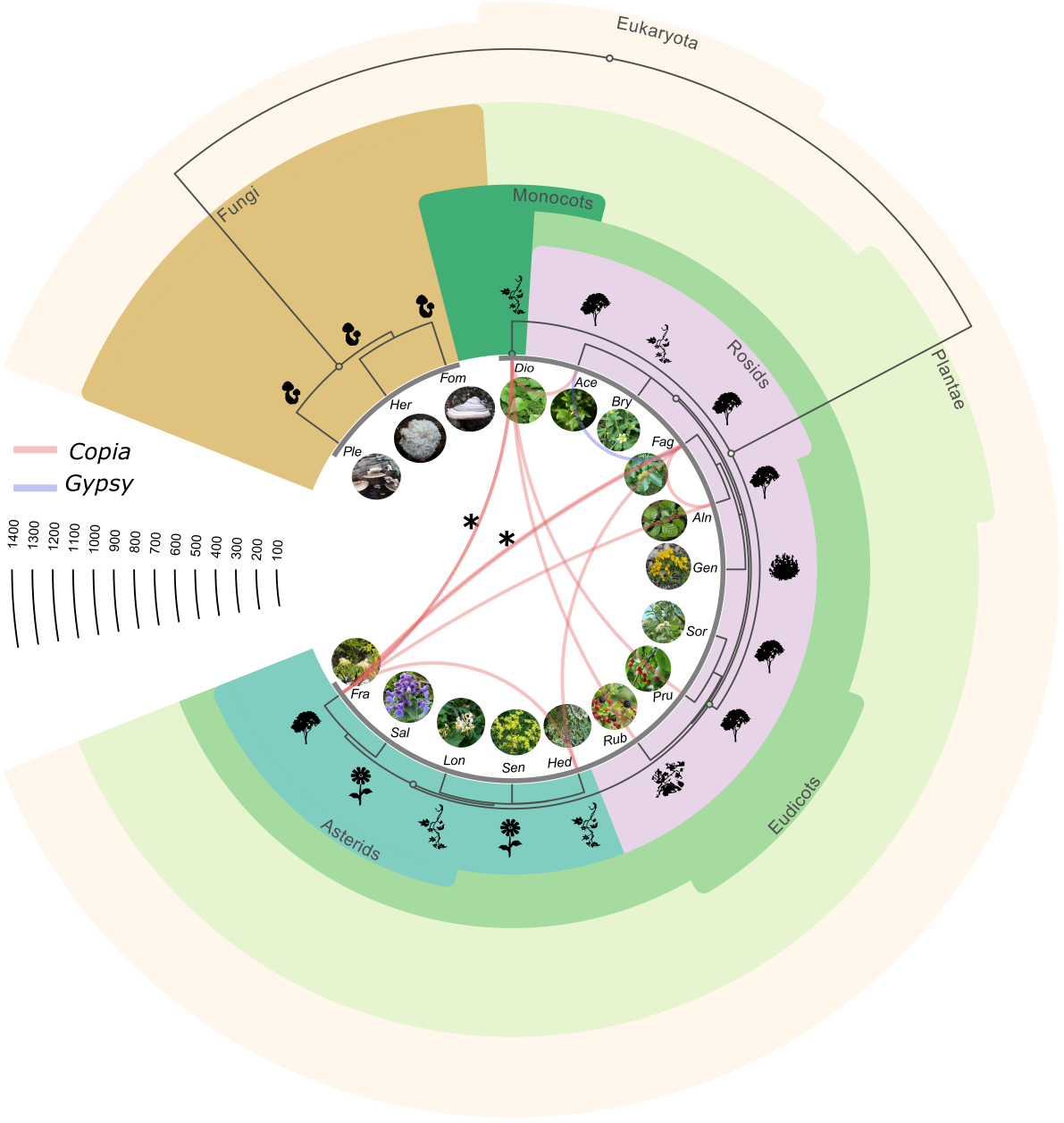
Species tree

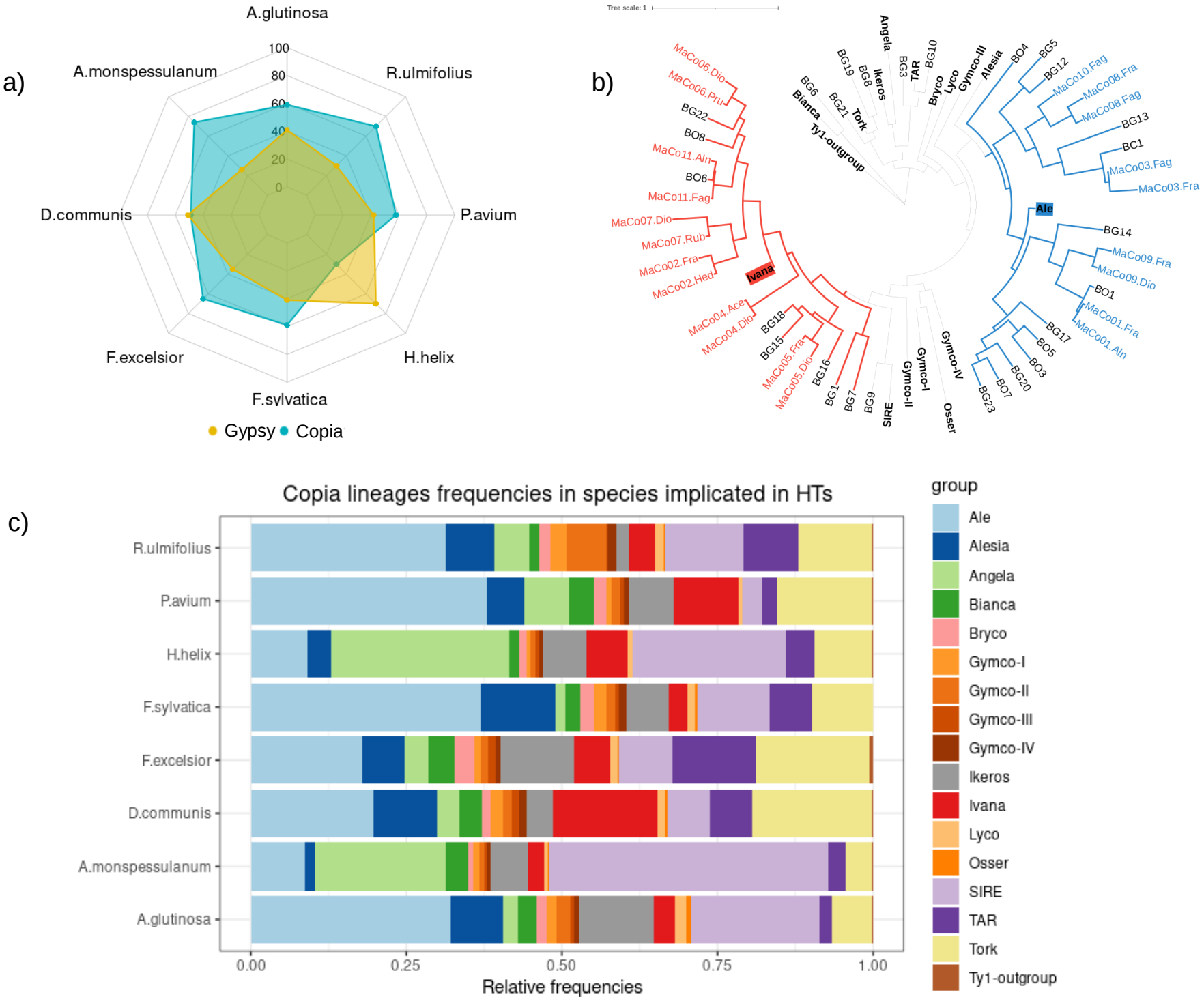
Step 10
PD test using all available plant genomes




a)
In silico simulation of 400 HTs per species pair
A. thaliana
O. sativa
B. distachyon
160 My
46 My
200 HTs
200 HTs
200 HTs
200 HTs
HTs identification using INTERCHANGE
b)
Category
False Positive
False Negative
True Positive
A. thaliana vs O. sativa
B. distachyon vs O. sativa
Genes
TEs
Numbers_of_HT_candidates
0
200
400
600
800
1000
1200
>0.2Kbp
>0.5Kbp
>1Kbp
>1.5Kbp
>2Kbp
Scaffold_size


100 MYA
108 MYA
117 MYA
Eudicots
160 MYA
Monocots
Prunus persica
Populus trichocarpa
Citrus clementina
Vitis vinifera
Phoenix dactylifera
BO7
HT4
HT5
BO3
BO4
HT3
BO2
BO1
BC1
HT1
HT2


Eukaryota
Fungi
Monocots
Plantae
Rosids
Eudicots
Asterids
Copia
Gypsy
1400
1300
1200
1100
1000
900
800
700
600
500
400
300
200
100
Dio
Ace
Bry
Fag
Aln
Gen
Sor
Pru
Rub
Hed
Sen
Lon
Sal
Fra
Ple
Her
Fom
*
*


a)
A.glutinosa
R.ulmifolius
P.avium
H.helix
F.sylvatica
F.excelsior
D.communis
A.monspessulanum
100
80
60
40
20
0
Gypsy
Copia
b)
Tree scale: 1
MaCo06.Dio
MaCo06.Pru
BG22
BO8
MaCo11.Aln
BO6
MaCo11.Fag
MaCo07.Dio
MaCo07.Rub
MaCo02.Fra
MaCo02.Hed
Ivana
MaCo04.Ace
MaCo04.Dio
BG18
BG15
MaCo05.Fra
MaCo05.Dio
BG16
BG1
BG7
BG9
SIRE
Gymco-I
Gymco-II
Gymco-IV
Osser
BG23
BO7
BG20
BO3
BO5
BG17
MaCo01.Aln
MaCo01.Fra
BO1
MaCo09.Dio
MaCo09.Fra
BG14
Ale
MaCo03.Fra
MaCo03.Fag
BC1
BG13
MaCo08.Fag
MaCo08.Fra
MaCo10.Fag
BG12
BG5
BO4
Alesia
Gymco-III
Lyco
Bryco
TAR
BG10
BG3
Angela
Ikeros
BG8
BG19
Tork
BG21
Bianca
BG6
Ty1-outgroup
c)
Copia lineages frequencies in species implicated in HTs
R.ulmifolius
P.avium
H.helix
F.sylvatica
F.excelsior
D.communis
A.monspessulanum
A.glutinosa
0.00
0.25
0.50
0.75
1.00
Relative frequencies
group
Ale
Alesia
Angela
Bianca
Bryco
Gymco-I
Gymco-II
Gymco-III
Gymco-IV
Ikeros
Ivana
Lyco
Osser
SIRE
TAR
Tork
Ty1-outgroup